

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 01:23:55 ; Search time 44 Seconds
(without alignments)
2333.316 Million cell updates/sec

Title: US-09-834-271A-27

Perfect score: 185
Sequence: 1 ggccttaaggccctgcaatc.....tgtataagaataaggagctc 185

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185	100.0	185	3	US-09-384-305-22
2	185	100.0	185	3	US-09-258-377-27
3	183.4	99.1	185	3	US-09-384-305-21
4	183.4	99.1	185	3	US-09-258-377-26
5	180.2	97.4	185	3	US-09-384-305-20
6	180.2	97.4	185	3	US-09-258-377-25
7	158	85.4	10216	2	US-08-875-154-1
8	157.2	85.0	2604	3	US-09-264-097-3
9	157.2	85.0	2604	4	US-09-537-168-5
10	154	83.2	162	1	US-08-434-255-24
11	154	83.2	162	1	US-08-459-967-24
12	154	83.2	162	1	US-08-460-327-24
13	154	83.2	162	1	US-08-459-871-24
14	154	83.2	2084	1	US-08-720-859-3
15	154	83.2	2084	1	US-08-459-610-3
16	154	83.2	2084	2	US-08-343-804-3
17	154	83.2	2084	2	US-08-687-399-3
18	154	83.2	2084	3	US-08-600-908A-3
19	154	83.2	2084	3	US-08-683-838A-3
20	154	83.2	2084	3	US-09-182-859-3
21	154	83.2	2084	3	US-09-170-670-13
22	154	83.2	2084	3	US-09-193-088-31
23	154	83.2	2084	3	US-09-183-412-58
24	154	83.2	2084	4	US-09-290-734-13
25	154	83.2	2084	4	US-09-672-459-3
26	154	83.2	2084	4	US-09-636-252A-3
27	154	83.2	2084	4	US-09-545-586-13

28	154	83.2	2084	4	US-10-186-042-3	Sequence 3, Appli
29	154	83.2	2084	4	US-09-769-864-58	Sequence 58, Appl
30	44	23.8	44	3	US-09-384-305-15	Sequence 15, Appl
31	44	23.8	44	3	US-09-258-377-13	Sequence 19, Appl
32	39	21.1	42	2	US-08-873-479-37	Sequence 37, Appl
33	39	21.1	42	2	US-09-031-442A-9	Sequence 9, Appl
34	39	21.1	42	2	US-08-972-661A-25	Sequence 25, Appl
35	39	21.1	42	3	US-09-258-377-9	Sequence 9, Appl
C 36	38.4	20.8	43	2	US-08-873-479-38	Sequence 38, Appl
C 37	38.4	20.8	43	2	US-09-031-442A-10	Sequence 10, Appl
C 38	38.4	20.8	43	2	US-08-972-661A-26	Sequence 26, Appl
C 39	38.4	20.8	43	3	US-09-258-377-10	Sequence 10, Appl
40	34	18.4	64081	4	US-09-790-988-1	Sequence 1, Appl
41	33.2	17.9	1511	1	US-07-991-867B-8	Sequence 8, Appl
42	33.2	17.9	1511	1	US-08-107-755A-8	Sequence 8, Appl
43	33.2	17.9	1511	2	US-08-544-332-8	Sequence 8, Appl
44	33.2	17.9	1511	4	US-09-370-861A-8	Sequence 8, Appl
C 45	32.6	17.6	8700	2	US-08-392-625-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-384-305-22
; Sequence 22, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:

; APPLICANT: Michael D. Thomas
; TITLE OF INVENTION: Polypeptides Having Pectin
; TITLE OF INVENTION: Acetylase Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5952.000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-384-305-22

Query Match 100.0%; Score 185; DB 3; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.3e-43;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGCTTAAGGGCTGCAATCGATTGTTTCAGAAAAGAGAACCATATAAAATACCTTGT	60
Db	1	GGCTTTAAGGGCTGCAATCGATTGTTTCAGAAAAGAGAACCATATAAAATACCTTGT	60
Qy	61	CTGTCATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGGCTGTGTAAAAAATAGGA	120
Db	61	CTGTCATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGGCTGTGTAAAAAATAGGA	120
Qy	121	ATAAAGGGGGTTCACATTATTTTACTGATATGATATATATTAATTGTATAGAAATGG	180
Db	121	ATAAAGGGGGTTCACATTATTTTACTGATATGATATATATTAATTGTATAGAAATGG	180
Qy	181	AGCTC 185	
Db	181	AGCTC 185	

RESULT 2
US-09-258-377-27
; Sequence 27, Application US/09258377
; Patent No. 6255075
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a

```
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/031,442
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-09-258-377-27

Query Match          100.0%; Score 185; DB 3; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.3e-43;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60

QY 61 CTGTCAATCAGACAGGGTATTTTATCTGTCCAGACTGTCGCTGTGTAATAATAGGA 120
DB 61 CTGTCAATCAGACAGGGTATTTTATCTGTCCAGACTGTCGCTGTGTAATAATAGGA 120

QY 121 ATAAAGGGGGTTCACATTTTACTGATATGTAATAATTTGTAAGAAAATGG 180
DB 121 ATAAAGGGGGTTCACATTTTACTGATATGTAATAATTTGTAAGAAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 3
US-09-384-305-21
; Sequence 21, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:
; APPLICANT: Michael D. Thomas
; APPLICANT: Kimberly M. Brown
; TITLE OF INVENTION: Polypeptides Having Pectin
; FILE REFERENCE: 5952.000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-384-305-21

Query Match          99.1%; Score 183.4; DB 3; Length 185;
Best Local Similarity 99.5%; Pred. No. 1.5e-42;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60

QY 61 CTGTCAATCAGACAGGGTATTTTATCTGTCCAGACTGTCGCTGTGTAATAATAGGA 120
DB 61 CTGTCAATCAGACAGGGTATTTTATCTGTCCAGACTGTCGCTGTGTAATAATAGGA 120

QY 121 ATAAAGGGGGTTCACATTTTACTGATATGTAATAATTTGTAAGAAAATGG 180
DB 121 ATAAAGGGGGTTCACATTTTACTGATATGTAATAATTTGTAAGAAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 4
US-09-384-305-20
; Sequence 20, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:
; APPLICANT: Michael D. Thomas
; APPLICANT: Kimberly M. Brown
; TITLE OF INVENTION: Polypeptides Having Pectin
; FILE REFERENCE: 5952.000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-384-305-20

Query Match          97.4%; Score 180.2; DB 3; Length 185;
Best Local Similarity 98.4%; Pred. No. 1.2e-41;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
```

```
DB 181 AGCTC 185

RESULT 4
US-09-258-377-26
; Sequence 26, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widheer, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/031,442
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-09-258-377-26

Query Match          99.1%; Score 183.4; DB 3; Length 185;
Best Local Similarity 99.5%; Pred. No. 1.5e-42;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60

QY 61 CTGTCAATCAGACAGGGTATTTTATCTGTCCAGACTGTCGCTGTGTAATAATAGGA 120
DB 61 CTGTCAATCAGACAGGGTATTTTATCTGTCCAGACTGTCGCTGTGTAATAATAGGA 120

QY 121 ATAAAGGGGGTTCACATTTTACTGATATGTAATAATTTGTAAGAAAATGG 180
DB 121 ATAAAGGGGGTTCACATTTTACTGATATGTAATAATTTGTAAGAAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 5
US-09-384-305-20
; Sequence 20, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:
; APPLICANT: Michael D. Thomas
; APPLICANT: Kimberly M. Brown
; TITLE OF INVENTION: Polypeptides Having Pectin
; FILE REFERENCE: 5952.000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-384-305-20

Query Match          97.4%; Score 180.2; DB 3; Length 185;
Best Local Similarity 98.4%; Pred. No. 1.2e-41;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
```

```
Db 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAAAGAACCATAAAAATACCTTGT 60
Qy 61 CTGTCTATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAATAAATAGGA 120
Db 61 CTGTCTATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAATAAATAGGA 120
Qy 121 ATAAAGGGGGTGTGACATATTTTACTGATATGTAATATATATATTTGTAAGAAATGG 180
Db 121 ATAAAGGGGGTGTGATTTATTTTACTGATATGTAATAATATATTTGTAAGAAATGG 180
Qy 181 AGCTC 185
Db 181 AGCTC 185

RESULT 6
US-09-258-377-25
; Sequence 25, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods for Producing A polypeptide in a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; EARLIER FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/031,442
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-09-258-377-25

Query Match 97.4%; Score 180.2; DB 3; Length 185;
Best Local Similarity 98.4%; Pred. No. 1.2e-41;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAAAGAACCATAAAAATACCTTGT 60
Db 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAAAGAACCATAAAAATACCTTGT 60
Qy 61 CTGTCTATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAATAAATAGGA 120
Db 61 CTGTCTATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAATAAATAGGA 120
Qy 121 ATAAAGGGGGTGTGACATATTTTACTGATATGTAATATATATTTGTAAGAAATGG 180
Db 121 ATAAAGGGGGTGTGATTTATTTTACTGATATGTAATAATATATTTGTAAGAAATGG 180
Qy 181 AGCTC 185
Db 181 AGCTC 185

RESULT 7
US-08-875-154-1
; Sequence 1, Application US/08875154
; Patent No. 5882888
; GENERAL INFORMATION:
; APPLICANT: Jorgensen, Streen Troels
; TITLE OF INVENTION: DNA Integration By Transporation
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5882888 No. 5882888disk of No. 5882888th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
```

```
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/875,154
FILING DATE: 17-JUL-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4381.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10216 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "PMOL553"
US-08-875-154-1

Query Match 85.4%; Score 158; DB 2; Length 10216;
Best Local Similarity 97.0%; Pred. No. 4.2e-35;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 14 TGAATCGATTGTTTGAGAAAGAAAGAACCATAAAAATACCTTGTGTCATCAGACA 73
Db 2237 TGAATCGATTGTTTGAGAAAGAAAGAACCATAAAAATACCTTGTGTCATCAGACA 2236
Qy 74 GGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAATAGGAATAAGGGGGTT 133
Db 2297 GGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAATAGGAATAAGGGGGTT 133
Qy 134 GACATTTTACTGATATGTAATAATATTTGTAAGAAATG 179
Db 2357 GTTATTTTACTGATATGTAATAATATTTGTAAGAAATG 2402

RESULT 8
US-09-264-097-3
; Sequence 3, Application US/09264097
; Patent No. 6287826
; GENERAL INFORMATION:
; APPLICANT: No. 6287826man, Barrie Edmund
; APPLICANT: Hendriksen, Hanne Vang
; TITLE OF INVENTION: Enzymatic Preparation of Glucose Syrup
; FILE REFERENCE: 5278.200-US
; CURRENT APPLICATION NUMBER: US/09/264,097
; CURRENT FILING DATE: 1999-03-08
; EARLIER APPLICATION NUMBER: PA 0321/98
; EARLIER FILING DATE: 1998-03-09
; EARLIER APPLICATION NUMBER: 60/079,209
; EARLIER FILING DATE: 1998-03-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2604
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
US-09-264-097-3

Query Match 85.0%; Score 157.2; DB 3; Length 2604;
Best Local Similarity 98.1%; Pred. No. 5.2e-35;
Matches 159; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 18 ATCGATTGTTTGAGAAAGAAAGAACCATAAAAATACCTTGTGTCATCAGACAGGGT 77
```

593 ATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGCTGTCATCAGACGGGT 652
78 ATTTTATCTGTCAGACTGTCGCTGTGTAAAAATAGGAATAAGAGGGGGTTGACA 137
653 ATTTTATCTGTCAGACTGTCGCTGTGTAAAAATAGGAATAAGAGGGGGTTGTTA 712
138 TTATTTACTGATATGATATATAATTTGTATAAGAAATG 179
713 TTATTTACTGATATGATATAATTTGTATAAGAAATG 754

RESULT 9

US-09-537-168-5
; Sequence 5, Application US/09537168
; Patent No. 6410295
; GENERAL INFORMATION:
; APPLICANT: Andersen, Carsten
; APPLICANT: Jorgensen, Christel T.
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Svendsen, Allan
; APPLICANT: Kjaerulf, Soren
; TITLE OF INVENTION: Alpha-Amylase Variants
; FILE REFERENCE: 5986.200-US
; CURRENT APPLICATION NUMBER: US/09/537,168
; EARLIER FILING DATE: 2000-03-29
; EARLIER APPLICATION NUMBER: PA 1999 00437
; EARLIER FILING DATE: 1999-03-30
; EARLIER APPLICATION NUMBER: 60/127,427
; EARLIER FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2604
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
US-09-537-168-5

Query Match 85.0%; Score 157.2; DB 4; Length 2604;
Best Local Similarity 98.1%; Pred. No. 5.2e-35;
Matches 159; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 18 ATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGCTGTCATCAGACGGGT 77
Db 593 ATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGCTGTCATCAGACGGGT 652
QY 78 ATTTTATCTGTCAGACTGTCGCTGTGTAAAAATAGGAATAAGAGGGGGTTGACA 137
Db 653 ATTTTATCTGTCAGACTGTCGCTGTGTAAAAATAGGAATAAGAGGGGGTTGTTA 712
QY 138 TTATTTACTGATATGATATAATTTGTATAAGAAATG 179
Db 713 TTATTTACTGATATGATATAATTTGTATAAGAAATG 754

RESULT 10

US-08-434-255-24
; Sequence 24, Application US/08434255
; Patent No. 5621089
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dambmann, Claus
; APPLICANT: Aaslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESSES:
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,255
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 3764.400-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-434-255-24

Query Match 83.2%; Score 154; DB 1; Length 162;
Best Local Similarity 96.9%; Pred. No. 2.3e-34;
Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 18 ATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGCTGTCATCAGACGGGT 77
Db 1 ATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGCTGTCATCAGACGGGT 60
QY 78 ATTTTATCTGTCAGACTGTCGCTGTGTAAAAATAGGAATAAGAGGGGGTTGACA 137
Db 61 ATTTTATCTGTCAGACTGTCGCTGTGTAAAAATAGGAATAAGAGGGGGTTGTTA 120
QY 138 TTATTTACTGATATGATATAATTTGTATAAGAAATG 179
Db 121 TTATTTACTGATATGATATAATTTGTATAAGAAATG 162

RESULT 11

US-08-459-967-24
; Sequence 24, Application US/08459967
; Patent No. 5622841
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dambmann, Claus
; APPLICANT: Aaslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESSES:
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,967
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/434,255
FILING DATE: 03-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.


```

; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 3764,400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-459-967-24

Query Match      83.2%; Score 154; DB 1; Length 162;
Best Local Similarity 96.9%; Pred. No. 2.3e-34;
Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTGTCTCATCAGACAGGTT 77
Db 1 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTGTCTCATCAGACAGGTT 60

QY 78 ATTTTATGCTGTCAGACTGTCGCTGTGTATAAATAAGGAGGAGGAGGAGGAGGAGGAGG 137
Db 61 ATTTTATGCTGTCAGACTGTCGCTGTGTATAAATAAGGAGGAGGAGGAGGAGGAGGAGG 120

QY 138 TTATTTTACTGATGATATATAATATATATATATATATATATATATATATATATATAT 179
Db 121 TTATTTTACTGATGATATATAATATATATATATATATATATATATATATATATATAT 162

RESULT 12
US-08-460-327-24
; Sequence 24, Application US/08460327
; Patent No. 5622850
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dambmann, Claus
; APPLICANT: Aaslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5622850 No. 5622850disk of No. 5622850th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,327
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/434,255
; FILING DATE: 03-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Agis Dr., Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 3764,400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; Query Match      83.2%; Score 154; DB 1; Length 162;
; Best Local Similarity 96.9%; Pred. No. 2.3e-34;
; Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTGTCTCATCAGACAGGTT 77
Db 1 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTGTCTCATCAGACAGGTT 60

QY 78 ATTTTATGCTGTCAGACTGTCGCTGTGTATAAATAAGGAGGAGGAGGAGGAGGAGGAGG 137
Db 61 ATTTTATGCTGTCAGACTGTCGCTGTGTATAAATAAGGAGGAGGAGGAGGAGGAGGAGG 120

QY 138 TTATTTTACTGATGATATATAATATATATATATATATATATATATATATATATATAT 179
Db 121 TTATTTTACTGATGATATATAATATATATATATATATATATATATATATATATATAT 162

RESULT 13
US-08-459-871-24
; Sequence 24, Application US/08459871
; Patent No. 5650326
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dambmann, Claus
; APPLICANT: Aaslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5650326 No. 5650326disk of No. 5650326th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,871
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/434,255
; FILING DATE: 03-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Agis Dr., Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 3764,400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; Query Match      83.2%; Score 154; DB 1; Length 162;
; Best Local Similarity 96.9%; Pred. No. 2.3e-34;
; Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTGTCTCATCAGACAGGTT 77
Db 1 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTGTCTCATCAGACAGGTT 60

QY 78 ATTTTATGCTGTCAGACTGTCGCTGTGTATAAATAAGGAGGAGGAGGAGGAGGAGGAGG 137
Db 61 ATTTTATGCTGTCAGACTGTCGCTGTGTATAAATAAGGAGGAGGAGGAGGAGGAGGAGG 120

QY 138 TTATTTTACTGATGATATATAATATATATATATATATATATATATATATATATATAT 179
Db 121 TTATTTTACTGATGATATATAATATATATATATATATATATATATATATATATATAT 162

RESULT 14
US-08-459-871-24
; Sequence 24, Application US/08459871
; Patent No. 5650326
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dambmann, Claus
; APPLICANT: Aaslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5650326 No. 5650326disk of No. 5650326th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,871
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/434,255
; FILING DATE: 03-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Agis Dr., Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 3764,400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; Query Match      83.2%; Score 154; DB 1; Length 162;
; Best Local Similarity 96.9%; Pred. No. 2.3e-34;
; Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTGTCTCATCAGACAGGTT 77
Db 1 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTGTCTCATCAGACAGGTT 60

QY 78 ATTTTATGCTGTCAGACTGTCGCTGTGTATAAATAAGGAGGAGGAGGAGGAGGAGGAGG 137
Db 61 ATTTTATGCTGTCAGACTGTCGCTGTGTATAAATAAGGAGGAGGAGGAGGAGGAGGAGG 120

QY 138 TTATTTTACTGATGATATATAATATATATATATATATATATATATATATATATATAT 179
Db 121 TTATTTTACTGATGATATATAATATATATATATATATATATATATATATATATATAT 162
```

US-08-460-327-24

Query Match 83.2%; Score 154; DB 1; Length 162;
Best Local Similarity 96.9%; Pred. No. 2.3e-34;
Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTGTCTCATCAGACAGGTT 77

Db 1 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTGTCTCATCAGACAGGTT 60

QY 78 ATTTTATGCTGTCAGACTGTCGCTGTGTATAAATAAGGAGGAGGAGGAGGAGGAGG 137

Db 61 ATTTTATGCTGTCAGACTGTCGCTGTGTATAAATAAGGAGGAGGAGGAGGAGGAGG 120

QY 138 TTATTTTACTGATGATATATAATATATATATATATATATATATATATATATATATAT 179

Db 121 TTATTTTACTGATGATATATAATATATATATATATATATATATATATATATATATAT 162

RESULT 13

US-08-459-871-24

; Sequence 24, Application US/08459871
; Patent No. 5650326
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dambmann, Claus
; APPLICANT: Aaslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5650326 No. 5650326disk of No. 5650326th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,871
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/434,255
; FILING DATE: 03-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Agis Dr., Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 3764,400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-459-871-24

Query Match 83.2%; Score 154; DB 1; Length 162;
Best Local Similarity 96.9%; Pred. No. 2.3e-34;
Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTGTCTCATCAGACAGGTT 77

Db 1 ATCGATTGTTTGAGAAAGAGAGACCATATAAATACCTTGTCTGTCTCATCAGACAGGTT 60

QY 78 ATTTTATGCTGTCAGACTGTCGCTGTGTATAAATAAGGAGGAGGAGGAGGAGGAGG 137

Db 61 ATTTTATGCTGCCAGACTGTCGCTGTGTATAAATAAGGAGGAGGTTGTTA 120
QY 138 TTATTTTACTGATATGATATAATAAATTTGTATAAGAAATG 179
Db 121 TTATTTTACTGATATGATAAATAATAATTTGTATAAGAAATG 162

RESULT 14

US-08-720-899-3
; Sequence 3, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: ANYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5753460 No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08720,899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2084 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 250..1794
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 250..342
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 343..1791
US-08-720-899-3

Query Match 83.2%; Score 154; DB 1; Length 2084;
Best Local Similarity 96.9%; Pred. No. 3.9e-34;
Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 18 ATCATGTTGTTGAGAAAAGAGAGACCATAAATACCTTGTCTGTCATCAGACAGGTT 77
Db 73 ATCGATTGTTTGAGAAAAGAGAGACCATAAATACCTTGTCTGTCATCAGACAGGTT 132
QY 78 ATTTTATGCTGCCAGACTGTCGCTGTGTATAAATAAGGAGGAGTTGACA 137

Db 133 ATTTTATGCTGCCAGACTGTCGCTGTGTATAAATAAGGAGGAGTTGTTA 192
QY 138 TTATTTTACTGATATGATATAATAAATTTGTATAAGAAATG 179
Db 193 TTATTTTACTGATATGATAAATAATAATTTGTATAAGAAATG 234

RESULT 15

US-08-459-610-3
; Sequence 3, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: ANYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5801043 No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08459,610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2084 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 250..1794
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 250..342
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 343..1791
US-08-459-610-3

Query Match 83.2%; Score 154; DB 1; Length 2084;
Best Local Similarity 96.9%; Pred. No. 3.9e-34;
Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 18 ATCATGTTGTTGAGAAAAGAGAGACCATAAATACCTTGTCTGTCATCAGACAGGTT 77
Db 73 ATCGATTGTTTGAGAAAAGAGAGACCATAAATACCTTGTCTGTCATCAGACAGGTT 132
QY 78 ATTTTATGCTGCCAGACTGTCGCTGTGTATAAATAAGGAGGAGTTGACA 137
Db 133 ATTTTATGCTGCCAGACTGTCGCTGTGTATAAATAAGGAGGAGTTGTTA 192

Qy 138 TTATTTTACTGATATGTATATATAATTTGTATTAAGAAATG 179
|||
Db 193 TTATTTTACTGATATGTATATATAATTTGTATTAAGAAATG 234
|||

Search completed: March 19, 2004, 03:17:58
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 01:21:15 ; Search time 1837 Seconds

(without alignments)
3007.351 Million cell updates/sec

Title: US-09-834-271A-27

Perfect score: 185

Sequence: 1 ggccttaaggcctgaatc.....tgtataagaaaatggagctc 185

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_esthum.*
4: em_esthum.*
5: em_esthum.*
6: em_esthum.*
7: em_esthum.*
8: em_esthum.*
9: gb_est1.*
10: gb_est2.*
11: gb_est3.*
12: gb_est4.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estfun.*
17: em_gss_hum.*
18: em_gss_hum.*
19: em_gss_hum.*
20: em_gss_vrt.*
21: em_gss_vrt.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_pro.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43.4	23.5	427	14	CA819917
2	39.8	21.5	1201	13	EX356191
3	39.4	21.3	1201	9	AL554700
4	38.8	21.0	735	29	CE808066

C	5	38.6	20.9	1201	13	EX446359
C	6	38.6	20.9	1304	28	CC317939
C	7	37.6	20.3	392	10	AW292419
C	8	37.6	20.3	461	9	AW001963
C	9	37.6	20.3	462	12	BQ000230
C	10	37.6	20.3	467	9	AI989386
C	11	37.6	20.3	475	9	AI697642
C	12	37.6	20.3	564	9	AI479617
C	13	37.6	20.3	699	28	BZ072552
C	14	37.6	20.3	700	9	AU253100
C	15	37.4	20.2	894	9	AL543011
C	16	37.4	20.2	1010	10	BE614304
C	17	37.4	20.2	1038	29	CNS01TL7
C	18	36.8	19.9	700	28	BH922535
C	19	36.6	19.8	403	9	AI499797
C	20	36.6	19.8	404	9	AI570619
C	21	36.6	19.8	708	29	EX240763
C	22	36.6	19.8	822	29	CNS03LLL
C	23	36.4	19.7	792	29	CNS02D0S
C	24	36.4	19.7	1068	29	CNS00ETV
C	25	36.4	19.7	1200	13	EX426107
C	26	36.4	19.7	1201	13	EX442065
C	27	36.2	19.6	928	29	CNS00DKY
C	28	36.2	19.6	565	12	BM153720
C	29	36.2	19.5	650	28	BZ361891
C	30	36.2	19.5	695	28	BZ374571
C	31	36.2	19.5	938	29	CG288637
C	32	36.2	19.5	997	29	CG288649
C	33	36.2	19.5	1201	13	EX385059
C	34	35.8	19.4	397	28	AZ170524
C	35	35.6	19.2	498	13	EX569556
C	36	35.6	19.2	504	13	EX564726
C	37	35.6	19.2	512	13	EX568933
C	38	35.6	19.2	534	12	BI371578
C	39	35.6	19.2	639	28	AZ367577
C	40	35.6	19.2	868	28	AZ535134
C	41	35.6	19.2	887	28	AZ546631
C	42	35.6	19.2	915	29	CG184658
C	43	35.6	19.2	995	29	CG184655
C	44	35.6	19.2	1201	13	EX419991
C	45	35.4	19.1	315	12	BG627035

ALIGNMENTS

RESULT	1	427 bp	mRNA	linear	EST	09-DEC-2002
CA819917	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
LOCUS	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
DEFINITION	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
ACCESSION	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
VERSION	CA819917.1	427 bp	mRNA	linear	EST	09-DEC-2002
KEYWORDS	CA819917.1	427 bp	mRNA	linear	EST	09-DEC-2002
SOURCE	CA819917.1	427 bp	mRNA	linear	EST	09-DEC-2002
ORGANISM	CA819917.1	427 bp	mRNA	linear	EST	09-DEC-2002
REFERENCE	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
AUTHORS	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
1	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
2	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
3	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
4	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
5	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
6	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
7	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
8	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
9	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
10	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
11	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
12	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
13	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
14	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
15	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
16	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
17	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
18	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
19	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
20	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
21	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
22	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
23	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
24	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
25	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
26	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
27	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
28	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002
29	CA819917	427 bp	mRNA	linear	EST	09-DEC-2002

CA819917 sa082a05.y1 Gm-cl048 Glycine max cdna clone SOYBEAN CLONE ID:
Gm-cl048-2314 5', mRNA sequence.
CA819917
CA819917.1 GI:26268854
EST.
Glycine max (soybean)
Glycine max.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 427)
Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Corryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact: ccu@resgen.com web site:
 www.resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 418.
 Location/Qualifiers

FEATURES

1. .427
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl048-2314"
 /tissue_type="whole seedlings of greenhouse grown plants"
 /dev_stage="1 week old"
 /lab_host="DH10B"
 /clone_lib="Gm-cl048"
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
 XhoI; The Clark NIL was constructed and seed was provided
 by Dr. J. Specht, University of Nebraska (Shoemaker and
 Specht, 1995). The cDNA library was constructed from mRNA
 isolated from whole seedlings of 1 week old greenhouse
 grown plants. Complementary DNA was synthesized from a
 using a primer consisting of a poly(dT) sequence with a
 XhoI restriction site and a 3' anchor. EcoRI adapters were
 ligated to the blunt-ended cDNA fragments followed by XhoI
 digestion. The cDNA fragments were directionally cloned
 into the EcoRI-XhoI restriction site of the pBluescript
 vector. The ligated cDNA fragments were transformed into
 DH10B host cells (GibcoBRL). The library was constructed
 in cooperation with Dr. Paul Keim's laboratory at Northern
 Arizona University."

ORIGIN

Query Match 23.5%; Score 43.4; DB 14; Length 427;
 Best Local Similarity 53.9%; Pred. No. 1.2; Indels 0; Gaps 0;
 Matches 89; Conservative 0; Mismatches 76;
 QY 19 TCGATTGTTGAGAAAAGAGAACCATATAAATACCTGTCTGTCATCAGACAGGTA 78
 Db 164 TTGCTTTTAAAGGAAAGGGAAGAAACCTTTTAACTTTTAACTTAACTC 223
 QY 79 TTTTATGCTGTCAGACTGCTCGCTGTGTAAATAATAGATAAAGGGGGTTCACAT 138
 Db 224 TTTTCCAAATCCCTTTTTCAGAGAGAAAGAAAGAAAGAGCCCTTTTAT 283
 QY 139 TATTCTGATGATGATAATATATTTGTATAAGAAATGGAGC 183
 Db 284 TTTTAAATTTTATTTCAATTTTATAGAAATGGACC 328

RESULT 2
 BX356191 1201 bp mRNA linear EST 05-MAY-2003
 LOCUS clone CS01009YC24 5-PRIME, mRNA sequence.
 DEFINITION

ACCESSION BX356191
 VERSION BX356191.1 GI:30365992
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM

REFERENCE
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 1433.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS01009BBI2QPI&cluster=1433.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS01009BBI2QPI.

FEATURES

source

1. .1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS01009YC24"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 21.5%; Score 39.8; DB 13; Length 1201;
 Best Local Similarity 45.8%; Pred. No. 6.9; Indels 0; Gaps 0;
 Matches 71; Conservative 16; Mismatches 58;
 QY 23 TTGTTTGAGAAAAGAGAACCATATAAATACCTGTCTGTCATCAGACAGGTAATTTT 82
 Db 892 TTTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 951
 QY 83 TTATGCTGTCAGACTGCTCGCTGTGTAAATAATAGATAAAGGGGGTTCACATATT 142
 Db 952 TTTTAT 1011
 QY 143 TTACTGATATGAT 177
 Db 1012 TTTTAT 1046

RESULT 3
 AL554700/1201 bp mRNA linear EST 31-MAY-2003
 LOCUS clone CS01085YH07 5-PRIME, mRNA sequence.
 DEFINITION

ACCESSION AL554700
 VERSION AL554700.2 GI:31276510
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM

REFERENCE
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 8225.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS01085CD04QPI&cluster=8225.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS01085CD04QPI.

FEATURES

source

1. .1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

```
/clone="CSODI085VH07"
/cisue_type="PLACENTA COT 25-NORMALIZED"
/cisue_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dn)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 21.3%; Score 39.4; DB 9; Length 1201;
Best Local Similarity 39.3%; Pred. No. 8.7;
Matches 59; Conservative 32; Mismatches 59; Indels 0; Gaps 0;

QY 17 AATCATGTTGTTGAGAAAGAGAGACCAATAAAATACCTGTCTGTCTCATCAGACAGGG 76
DB 1010 AAAAAAATTTTAAATAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 951
QY 77 TATTTTATGCTGTCAGACATGCTCGCTGTGTGTAATAAATAAGTAATAAGGGGGTTCAC 136
DB 950 MAWWTTTTTTTTTTTTTTTTTTTTTTTTTVVVVVVVRRRRRAAAARDDGGGKKKT 891
QY 137 ATTATTTTACGATGATGATATATATTAATTT 166
DB 890 NYTTTTTTTTTTTTTTTAAATGATCAACTT 861

RESULT 4
CEB08066 735 bp DNA linear GSS 30-SEP-2003
LOCUS tigr-gss-dog-17000317853400 Dog Library Canis familiaris genomic,
DEFINITION Genomic survey sequence.
ACCESSION CEB08066
VERSION CEB08066.1 GI:37149050
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
REFERENCE 1 (bases 1 to 735)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
Kirkness, E.F., Baina, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Dalcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
Location/Qualifiers
1. .735
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/cisue_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match 21.0%; Score 38.8; DB 29; Length 735;
Best Local Similarity 53.2%; Pred. No. 14;
Matches 82; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 30 AGAAAAGAGAGACCAATAAAATACCTGTCTGTCTCATCAGACAGGGTATTTTATGCT 89
DB 15 ACAAAACAAACAAACAAACAAATTCCTGTGTTTTCCTCTTTTCTTTCTTTTCTTTT 74

/clone="CSODI085VH07"
/cisue_type="PLACENTA COT 25-NORMALIZED"
/cisue_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dn)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 20.9%; Score 38.6; DB 13; Length 1201;
Best Local Similarity 39.9%; Pred. No. 14;
Matches 69; Conservative 27; Mismatches 77; Indels 0; Gaps 0;

QY 5 TTAAGGGCTGCAATCGATTGTTTGAAGAAAGAGACCAATAAAATACCTGTCTGT 64
DB 1024 TGAABAWYTCGGTGTGGGTGTGGGGGAGAAAAACCCSYSMCTGAACWCCMCACWNA 965
QY 65 CATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCGCTGTGTAAAAAATAGGAATAA 124
DB 964 CACWGAAGACGATGGKTTGTGCCCTGTGGCGWYTKTTTTTCTGKRAAAAAAARAWW 905
QY 125 AGGGGGGTGACATTATTTTACTGATATGTAATAATAATTTGTATAAGAAAA 177
DB 904 CTGGGGGYCTGYTYYTITGTAFAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 852

RESULT 6
CC317939/c 1304 bp DNA linear GSS 14-MAY-2003
LOCUS CC317939
DEFINITION TAM32-25P7_EC1.1 TAM32 Gallus genomic clone TAM32-25P7,
```

```

tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution.
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/INL at:
www.bio.lnlnl.gov/bbrp/image/image.html The following repetitive
elements were found in this cDNA sequence: 1-30,
>AT rich#Low complexity
Seq primer: M13 Forward
PDVAs=Yes.

Location/Qualifiers
1. 392
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/db_xref="IMAGE:2725886"
/lab_host="DHI0B (Life Technologies)"
/clone_libs="NCI CGAP Sub4"
/note="Vector: pVT73b-Pac (Pharmacia) with a modified
polylinker. Site 1: Not 1; Site 2: Eco RI; The
NCI CGAP Sub4 library is a subtracted library derived from
I.M.A.G.E. Consortium/INL at: www.bio.lnlnl.gov/bbrp/image/image.html The following repetitive
elements were found in this cDNA sequence: 1-30,
>AT rich#Low complexity
Seq primer: M13 Forward
PDVAs=Yes.
the NCI-CGAP Sub2 library which is a subtracted library
derived from the NCI-CGAP Sub1 library, which is a
subtracted library derived or subtracted NCI-CGAP
mixture of 21 normalized or subtracted NCI-CGAP
libraries: NCI CGAP Co4, NCI CGAP Pr22, NCI CGAP Pr28,
NCI CGAP Co10, NCI CGAP Co16, NCI-CGAP Kids,
NCI CGAP Kid12, NCI CGAP Kid3, NCI CGAP Kid1,
NCI CGAP Kid2, NCI CGAP Br2, NCI CGAP Co8, NCI-CGAP_C11L1,
NCI CGAP_LyM2, NCI CGAP Brn23, NCI CGAP_Lu5,
NCI CGAP_Lu24, NCI CGAP_Lu19, NCI CGAP GC4, NCI CGAP_GC6,
NCI CGAP_Brn25. These 21 libraries were pooled and a
single-stranded DNA preparation of the resulting mixture
was used as a tracer in a subtractive hybridization with
a driver whose composition is detailed below:
NCI CGAP Kid3 pool 1 : L1AM 334-3337, 3682-3683,
3795-3803 (IMAGE CloneIDs 1323376-1323911,
1456008-1456775, 1500552-1502855) NCI CGAP Kids pool 1 :
L1AM 3338-3344, 3722-3725, 3778-3778 (IMAGE CloneIDs
1323912-1325831, 1471368-1472903, 1492104-1493255)
NCI CGAP_Lu5 pool 1 : L1AM 3575-3582, 3851-3854 (IMAGE
CloneIDs 1414920-1417991, 1520904-1522439) NCI CGAP GC4
pool 1 : L1AM 3164-3167, 3716-3720, 3733-3735 (IMAGE
CloneIDs 1257096-1258631, 1469064-1470983,
1475592-1476931) NCI CGAP Pr22 pool 1 : L1AM 2457-2459,
2753-2759, 3062-3068 (IMAGE CloneIDs 985608-986759,
110192-110195, 1217928-1220615) NCI CGAP Co10 pool 1 :
L1AM 2644-2653, 2871-2872 (IMAGE CloneIDs 1057416-1061255,
114584-1145351) Subtraction was performed as previously
described [Bonaldo, Lennon & Soares (1996): Normalization
and Subtraction: Two Approaches To Facilitate Gene
Discovery. Genome Research 6, 791-806.]
TAG TISSUE=lung
TAG LIB=NCI CGAP_Lu5
TAG_SEQ=CACAC"

th 20.3%; Score 37.6; DB 10; Length 392;
Similarity 72.1%; Pred. No. 33;
49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

1.1 AAAAATGAGTAATAGGGGGTGCATTATTTACTGATGTATAATATTGTAT 170
3.3 AAGAGAGAGGAGAACTAAGGTGTCTTAATATGACAGATAGTAAAAATATTGCTTT 252

1.1 AAGAAAT 178
|||||
3.3 AAGAAAT 260

AW001963 461 bp mRNA linear EST 09-MAR-2000
wt52d01.x1 Soares.thymus.NHFT Homo sapiens cDNA clone
IMAGE:2513953 3', mRNA sequence.

```

ACCESSION AW001963
 VERSION AW001963.1 GI:5848879
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 461)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapsb-remail.nih.gov
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 696 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 448.
 Location/Qualifiers
 1..461
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2513953"
 /dev_stage="fetal"
 /tissue="lung"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Lung; Vector: pTT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTCACCAATCGAGTGGAGCGGCCCAACGTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

Query Match 20.3%; Score 37.6; DB 9; Length 461;
 Best Local Similarity 72.1%; Pred. No. 31;
 Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 111 AAAAATAGGAATAAAGGGGGTTGACATTATTTTCTGATATGATAATAAATTGTAT 170
 Db 176 AAGACAGAGGAGGAACCTAAGGTTCATTAAATGACAGATAGTAATAAATTGCTTT 235

QY 171 AAGAAAAAT 178
 Db 236 AAGAAAAAT 243

RESULT 9
 BQ000230/c
 LOCUS BQ000230 462 bp mRNA linear EST 17-JUN-2002
 DEFINITION UI-H-DIO-avp-j-04-0-UI.61 NCI-CGAP_DIO Homo sapiens cDNA clone IMAGE:5882595 3', mRNA sequence.

ACCESSION BQ000230
 VERSION BQ000230.1 GI:19725130
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 462)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapsb-remail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende

cdNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 The following repetitive elements were found in this cDNA sequence: 433-454, SAT rich#Low_complexity (matched complement)
 Seq primer: ML3 FORWARD
 POLYA=Yes.

Location/Qualifiers
 1..462
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5882595"
 /tissue_type="Lung Focal Fibrosis"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP DIO"
 /note="Organ: Lung; Vector: pTT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI-CGAP DIO is a cDNA library containing the following tissue(s): A pool of Lung Focal Fibrosis. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pTT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ATACGGCGTC.
 TAG TISSUE=Lung with fibrosis
 TAG LIB=UI-H-DIO
 TAG_SEQ=ATACGGCGTC"

ORIGIN
 Query Match 20.3%; Score 37.6; DB 12; Length 462;
 Best Local Similarity 72.1%; Pred. No. 31;
 Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 111 AAAAATAGGAATAAAGGGGGTTGACATTATTTTCTGATATGATAATAAATTGTAT 170
 Db 270 AAGACAGAGGAGGAACCTAAGGTTCATTAAATGACAGATAGTAATAAATTGCTTT 211

QY 171 AAGAAAAAT 178
 Db 210 AAGAAAAAT 203

RESULT 10
 AI989386
 LOCUS AI989386 467 bp mRNA linear EST 08-SEP-1999
 DEFINITION wt-80e12.x1 Soares thymus NHFT Homo sapiens cDNA clone IMAGE:2513806 3', mRNA sequence.

ACCESSION AI989386
 VERSION AI989386.1 GI:5836309
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 467)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapsb-remail.nih.gov
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40UP from Gibco


```

FEATURES
  source
    High quality sequence stop: 442.
    Location/Qualifiers
      1. .467
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:2513806"
        /dev_stage="fetal"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="Soares.Chymus_NHFin"
        /notes="Organ: thymus, pooled; Vector: pT7T3D-Pac
        (Pharmacia) with a modified polylinker; Site 1: Not I;
        Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
        oligo(dT) primer [5',
        TGTTACCATCTGAGTGGAGCGCGCGCAAGCTTTTCTTTT 3'],
        double-stranded cDNA was ligated to Eco RI adaptors
        (Pharmacia), digested with Not I and cloned into the Not I
        and Eco RI sites of the modified pT7T3 vector. Library
        went through one round of normalization. Library
        constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
  Query Match      20.3%; Score 37.6; DB 9; Length 467;
  Best Local Similarity 72.1%; Pred. No. 31;
  Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 111 AAAAATAGGAATAAGGGGGTTCACATTATTTTACTGATATGATATATATATTTGAT 170
    |||||
Db 176 AAGAGAGGAGGAACTAAGTTCATTAAATGACAGATAAGTAAAAAATAATGCTTT 235
    |||||

QY 171 AAGAAAT 178
    |||||
Db 236 AAGAAAT 243

RESULT 11
AI697642 475 bp mRNA linear EST 18-DEC-1999
LOCUS we15902.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2341202 3',
DEFINITION mRNA sequence.
ACCESSION AI697642
VERSION AI697642
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 475)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 701 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 465.
Location/Qualifiers
  1. .475
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:2341202"
    /tissue_type="carcinoid"
    /lab_host="DH10B"

FEATURES
  source
    High quality sequence stop: 442.
    Location/Qualifiers
      1. .467
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:2513806"
        /dev_stage="fetal"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="Soares.Chymus_NHFin"
        /notes="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
        modified polylinker; Plasmid DNA from the normalized
        library NCI_CGAP_Lu5 was prepared, and ss circles were
        made in vitro. Following HAP purification, this DNA was
        used as tracer in a subtractive hybridization reaction.
        The driver was PCR-amplified cDNAs from a pool of 5,000
        clones made from the same library (clonoids
        141920-141791 and 152094-1522439). Subtraction by Bento
        Soares and M. Fatima Bonaldo. "

ORIGIN
  Query Match      20.3%; Score 37.6; DB 9; Length 475;
  Best Local Similarity 72.1%; Pred. No. 31;
  Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 111 AAAAATAGGAATAAGGGGGTTCACATTATTTTACTGATATGATATATATTTGAT 170
    |||||
Db 179 AAGAGAGGAGGAACTAAGTTCATTAAATGACAGATAAGTAAAAAATAATGCTTT 238
    |||||

QY 171 AAGAAAT 178
    |||||
Db 239 AAGAAAT 246

RESULT 12
AI479617 564 bp mRNA linear EST 14-APR-1999
LOCUS tm32b12.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:2158271 3',
DEFINITION mRNA sequence.
ACCESSION AI479617
VERSION AI479617
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 564)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
M.D., Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 685 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 446.
Location/Qualifiers
  1. .564
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:2158271"
    /tissue_type="B-cell, chronic lymphocytic leukemia"
    /lab_host="DH10B"
    /clone_lib="NCI_CGAP_CLL1"
    /notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
    polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
    was primed with a Not I - oligo(dT) primer [5',
    TGTTACCATCTGAGTGGAGCGCGCGCATGCTTTTCTTTT 3'],
    double-stranded cDNA was ligated to Eco RI
    adaptors (Pharmacia), digested with Not I and cloned into
    the Not I and Eco RI sites of the modified pT7T3 vector.
    Library is normalized, and was constructed by Bento
    Soares and M. Fatima Bonaldo. "

```


Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9876.f. For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE013DF11QP1&cluster=9876.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope Sequence ID : CS0DE013DF11QP1.

FEATURES

Location/Qualifiers
1. .894
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013YL22"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="vector: PCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match 20.2%; Score 37.4; DB 9; Length 894;
Best Local Similarity 42.3%; Pred. No. 29; Mismatches 64; Indels 0; Gaps 0;
Matches 63; Conservative 22;
QY 30 AGAAAGAGAGAGACCATATAACCTTCTGTCATCAGACAGGGTATTTTATGCT 89
Db 648 AAAAAAAAAAAAAAAAAAWTTTTTTTHMWAWTTTTTTTTTTTTTTTTTKK 707
QY 90 GTCAGACTGCCCTGTGTAAATAAGAAATAGAGGAGGGGTGACATTTTACTGA 149
Db 708 KKIDCATTTTWWWWWWWTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 767
QY 150 TATGTATAATATATTTGTATAGAAAAT 178
Db 768 TTTATATTWTAAKAAAAAAAAAATAAAW 796

Search completed: March 19, 2004, 03:16:13
Job time : 1838 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 02:14:41 ; Search time 745.5 Seconds
(without alignments)
917.966 Million cell updates/sec

Title: US-09-834-271A-27
Perfect score: 185
Sequence: 1 ggccttaaggccgcgaatc.....tgtatagaataaggagctc 185

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications NA:
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
 - 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq*
 - 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
 - 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
 - 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
 - 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq*
 - 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq*
 - 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
 - 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
 - 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
 - 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq*
 - 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq*
 - 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185	100.0	185	10	US-09-834-271A-27
2	185	100.0	185	14	US-10-406-025-6
3	183.4	99.1	185	10	US-09-834-271A-26
4	183.4	99.1	185	14	US-10-406-025-5
5	180.2	97.4	185	10	US-09-834-271A-25
6	180.2	97.4	185	14	US-10-406-025-7
7	158	85.4	2166	12	US-10-344-231-17
8	158	85.4	2166	12	US-10-363-332A-17
9	158	85.4	2267	12	US-10-344-231-18
10	158	85.4	2267	12	US-10-363-332A-18
11	158	85.4	2588	12	US-10-344-231-20
12	158	85.4	2588	12	US-10-363-332A-20
13	157.2	85.0	2604	14	US-10-146-327-5
14	154	83.2	2084	9	US-09-769-864-58
15	154	83.2	2084	9	US-09-854-346-9

16	154	83.2	2084	9	US-09-918-543-9	Sequence 9, Appli
17	154	83.2	2084	10	US-09-925-576C-9	Sequence 9, Appli
18	154	83.2	2084	12	US-10-665-667-58	Sequence 58, Appl
19	154	83.2	2084	12	US-10-665-667-58	Sequence 3, Appli
20	154	83.2	2084	14	US-10-184-771-3	Sequence 3, Appli
21	154	83.2	2084	14	US-10-184-771-3	Sequence 3, Appli
22	93.6	50.6	13222	10	US-09-736-116-60	Sequence 60, Appl
23	90.2	48.8	6837	10	US-09-928-847B-49	Sequence 49, Appl
24	44	23.8	44	10	US-09-834-271A-19	Sequence 19, Appl
25	39	21.1	42	10	US-09-834-271A-9	Sequence 9, Appli
26	38.4	20.8	43	10	US-09-834-271A-10	Sequence 10, Appl
27	37.6	20.3	2141	15	US-10-104-047-1737	Sequence 1737, Ap
28	36.4	19.7	3673778	14	US-10-312-841-2	Sequence 221, Appl
29	35.8	19.4	9121	9	US-09-070-927A-221	Sequence 7881, Ap
30	35.6	19.2	179	9	US-09-783-590-7881	Sequence 15736, A
31	35.2	19.0	512	15	US-10-027-632-15736	Sequence 145937,
32	35.2	19.0	827	15	US-10-027-632-145937	Sequence 389, App
33	35.2	19.0	19380	12	US-10-221-613-399	Sequence 577, App
34	35	18.9	11155	14	US-10-311-455-577	Sequence 198, App
35	35	18.9	14924	14	US-10-311-455-198	Sequence 22, Appl
36	35	18.9	14924	14	US-10-240-452-22	Sequence 1, Appli
37	34.8	18.8	3673778	14	US-10-312-841-1	Sequence 15549, A
38	34.6	18.7	2286	12	US-10-282-122A-15549	Sequence 470, App
39	34.6	18.7	14147	12	US-10-221-714A-470	Sequence 52, Appl
40	34.6	18.7	14147	14	US-10-172-086-52	Sequence 399, App
41	34.4	18.6	28588	9	US-09-764-887-399	Sequence 399, App
42	34.4	18.6	28588	14	US-10-073-961-399	Sequence 47, Appl
43	34.4	18.6	182328	12	US-10-235-192A-47	Sequence 18616, A
44	34	18.4	510	12	US-10-085-783A-18616	Sequence 18616, A
45	34	18.4	510	15	US-10-242-535A-18616	

ALIGNMENTS

RESULT 1
US-09-834-271A-27
; Sequence 27, Application US/09834271A
; Publication NO. US20030170876A1
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/834,271A
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US/09/258,377
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/031,442
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-09-834-271A-27

Query Match	100.0%	Score	185;	DB	10;	Length	185;
Best Local Similarity	100.0%	Pred. No.	1.1e-41;				
Matches	185;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	GGCTTAAAGGCGCTGCAATCGATTGTTTGAGAAAAGAAAGAACCAACCAATAAAATACCTTGT	60				
Db	1	GGCTTAAAGGCGCTGCAATCGATTGTTTGAGAAAAGAAAGAACCAACCAATAAAATACCTTGT	60				
Qy	61	CTGTTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGCCCTGTGTAATAAATAGGA	120				
Db	61	CTGTTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGCCCTGTGTAATAAATAGGA	120				
Qy	121	ATAAAGGGGGGTTTGACATTATTTTACTGATATGATATAATAATTTTGTATGAAGAAATGG	180				

```
Db 121 ATAAAGGGGGCTGACATTAATTTACTGATATGATATATATATTTGTTATAGAAATGG 180
Qy 181 AGCTC 185
Db 181 AGCTC 185

RESULT 2
US-10-406-025-6
; Sequence 6, Application US/10406025
; Publication No. US20030186380A1
; GENERAL INFORMATION:
; APPLICANT: Thomas, Michael D.
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Methods for producing secreted polypeptides having L-asparaginase
; FILE REFERENCE: 10289.200-US
; CURRENT APPLICATION NUMBER: US/10/406,025
; PRIOR FILING DATE: 2003-04-01
; PRIOR FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-10-406-025-6

Query Match 100.0%; Score 185; DB 14; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.1e-41; Indels 0; Gaps 0;
Matches 185; Conservative 0; Mismatches 0;

Qy 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAAGAGAGACCATAAAAAATACCTTGT 60
Db 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAAGAGAGACCATAAAAAATACCTTGT 60
Qy 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCACACTGTCGGCTGTGTAATAAATAGGA 120
Db 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCACACTGTCGGCTGTGTAATAAATAGGA 120
Qy 121 ATAAAGGGGGTTCACATTAATTTTACTGATATGATATATATATTTGTTATAGAAAAATGG 180
Db 121 ATAAAGGGGGTTCACATTAATTTTACTGATATGATATATATATTTGTTATAGAAAAATGG 180
Qy 181 AGCTC 185
Db 181 AGCTC 185

RESULT 3
US-09-834-271A-26
; Sequence 26, Application US/09834271A
; Publication No. US20030170876A1
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/834,271A
; CURRENT FILING DATE: 2001-04-12
; PRIOR FILING DATE: 1999-02-26
; PRIOR FILING DATE: 1999-02-26
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/031,442
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 185
```

```
; TYPE: DNA
; ORGANISM: Bacillus
US-09-834-271A-26

Query Match 99.1%; Score 183.4; DB 10; Length 185;
Best Local Similarity 99.5%; Pred. No. 2.9e-41; Indels 0; Gaps 0;
Matches 184; Conservative 0; Mismatches 1;

Qy 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAAGAGAGACCATAAAAAATACCTTGT 60
Db 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAAGAGAGACCATAAAAAATACCTTGT 60
Qy 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCACACTGTCGGCTGTGTAATAAATAGGA 120
Db 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCACACTGTCGGCTGTGTAATAAATAGGA 120
Qy 121 ATAAAGGGGGTTCACATTAATTTTACTGATATGATATATATATTTGTTATAGAAAAATGG 180
Db 121 ATAAAGGGGGTTCACATTAATTTTACTGATATGATATATATATTTGTTATAGAAAAATGG 180
Qy 181 AGCTC 185
Db 181 AGCTC 185

RESULT 4
US-10-406-025-5
; Sequence 5, Application US/10406025
; Publication No. US20030186380A1
; GENERAL INFORMATION:
; APPLICANT: Thomas, Michael D.
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Methods for producing secreted polypeptides having L-asparaginase
; FILE REFERENCE: 10289.200-US
; CURRENT APPLICATION NUMBER: US/10/406,025
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/369,192
; PRIOR FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-10-406-025-5

Query Match 99.1%; Score 183.4; DB 14; Length 185;
Best Local Similarity 99.5%; Pred. No. 2.9e-41; Indels 0; Gaps 0;
Matches 184; Conservative 0; Mismatches 1;

Qy 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAAGAGAGACCATAAAAAATACCTTGT 60
Db 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAAGAGAGACCATAAAAAATACCTTGT 60
Qy 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCACACTGTCGGCTGTGTAATAAATAGGA 120
Db 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCACACTGTCGGCTGTGTAATAAATAGGA 120
Qy 121 ATAAAGGGGGTTCACATTAATTTTACTGATATGATATATATATTTGTTATAGAAAAATGG 180
Db 121 ATAAAGGGGGTTCACATTAATTTTACTGATATGATATATATATTTGTTATAGAAAAATGG 180
Qy 181 AGCTC 185
Db 181 AGCTC 185

RESULT 5
US-09-834-271A-25
; Sequence 25, Application US/09834271A
; Publication No. US20030170876A1
```

GENERAL INFORMATION:
APPLICANT: Widner, William
APPLICANT: Sloma, Alan
APPLICANT: Thomas, Michael D.
TITLE OF INVENTION: Methods For Producing A polypeptide In a
TITLE OF INVENTION: Bacillus Cell
FILE REFERENCE: 5455-200-US
CURRENT APPLICATION NUMBER: US/09/834,271A
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: US/09/258,377
PRIOR FILING DATE: 1999-02-26
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/031.442
PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-26
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 25
LENGTH: 185
TYPE: DNA
ORGANISM: Bacillus
US-09-834-271A-25

Query Match 97.4%; Score 180.2; DB 10; Length 185;
Best Local Similarity 98.4%; Pred. No. 2.3e-40;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTCAATCGATTGTTGAGAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTCAATCGATTGTTGAGAAAGAGAGACCATATAAAATACCTTGT 60

QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGCCGTGTGTAATAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGCCGTGTGTAATAATAGGA 120

QY 121 ATAAAGGGGGTGGACATTTTACTGATGATGATATAATAATTTGTAAGAAATGG 180
DB 121 ATAAAGGGGGTGGTATTTTACTGATGATGATATAATAATTTGTAAGAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 6
US-10-406-025-7
Sequence 7, Application US/10406025
Publication No. US20030186380A1
GENERAL INFORMATION:
APPLICANT: No. US20030186380A1;zymes Biotech, Inc.
APPLICANT: Thomas, Michael D.
APPLICANT: Sloma, Alan
TITLE OF INVENTION: Methods for producing secreted polypeptides having L-asparaginase
TITLE OF INVENTION: activity
FILE REFERENCE: 10289.200-US
CURRENT APPLICATION NUMBER: US/10/406,025
CURRENT FILING DATE: 2003-04-01
PRIOR APPLICATION NUMBER: US 60/369,192
PRIOR FILING DATE: 2002-04-01
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.2
SEQ ID NO 7
LENGTH: 185
TYPE: DNA
ORGANISM: Bacillus
US-10-406-025-7

Query Match 97.4%; Score 180.2; DB 14; Length 185;
Best Local Similarity 98.4%; Pred. No. 2.3e-40;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTCAATCGATTGTTGAGAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTCAATCGATTGTTGAGAAAGAGAGACCATATAAAATACCTTGT 60

QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGTCGCTGTGTAAAAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGTCGCTGTGTAAAAATAGGA 120

QY 121 ATAAAGGGGGTGGACATTTTACTGATGATGATATAATAATTTGTAAGAAATGG 180
DB 121 ATAAAGGGGGTGGTATTTTACTGATGATGATATAATAATTTGTAAGAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 7
US-10-344-231-17
Sequence 17, Application US/10344231
Publication No. US20040038845A1
GENERAL INFORMATION:
APPLICANT: Pedersen, Poul
APPLICANT: Roggen, Erwin Ludo
TITLE OF INVENTION: Improved production of proteases with inhibitors
FILE REFERENCE: 10064.204-US
CURRENT APPLICATION NUMBER: US/10/344,231
CURRENT FILING DATE: 2003-02-10
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.2
SEQ ID NO 17
LENGTH: 2166
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-10-344-231-17

Query Match 85.4%; Score 158; DB 12; Length 2166;
Best Local Similarity 97.0%; Pred. No. 8.4e-34;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTGAGAAAGAGAGACCATATAAAATACCTTGTGTATCAGACA 73
DB 172 TGCATCGATTGTTGAGAAAGAGAGACCATATAAAATACCTTGTGTATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGCCGTGTGTAAAAATAGGAATAGGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGCCGTGTGTAAAAATAGGAATAGGGGGGTT 291

QY 134 GACATTTTACTGATGATGATATAATAATTTGTAAGAAATG 179
DB 292 GTTATTTTACTGATGATGATATAATAATTTGTAAGAAATG 337

RESULT 8
US-10-363-332A-17
Sequence 17, Application US/10363332A
Publication No. US20040038375A1
GENERAL INFORMATION:
APPLICANT: Pedersen, Poul
APPLICANT: No. US20040038375A1;regaard-Madsen, Mads
TITLE OF INVENTION: Method For Screening Highly Active Proteases and Inhibitors
FILE REFERENCE: 10080.204-US
CURRENT APPLICATION NUMBER: US/10/363,332A
CURRENT FILING DATE: 2003-05-27
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.2
SEQ ID NO 17
LENGTH: 2166
TYPE: DNA
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Construct A
US-10-363-332A-17

Query Match 85.4%; Score 158; DB 12; Length 2166;

Best Local Similarity 97.0%; Pred. No. 8.4e-34; DB 12; Length 2267;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 291

QY 134 GACATTATTTTACTGATATGATATAATAATTTGTATAAGAAAATG 179
DB 292 GTTATTTTACTGATATGATATAATAATTTGTATAAGAAAATG 337

RESULT 9
US-10-344-231-18
; Sequence 18, Application US/10344231
; Publication No. US20040038845A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; APPLICANT: Roggen, Erwin Ludo
; TITLE OF INVENTION: Improved production of proteases with inhibitors
; FILE REFERENCE: 10064.204-US
; CURRENT APPLICATION NUMBER: US/10/344,231
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 2267
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-344-231-18

Query Match 85.4%; Score 158; DB 12; Length 2267;
Best Local Similarity 97.0%; Pred. No. 8.5e-34;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 291

QY 134 GACATTATTTTACTGATATGATATAATAATTTGTATAAGAAAATG 179
DB 292 GTTATTTTACTGATATGATATAATAATTTGTATAAGAAAATG 337

RESULT 10
US-10-363-332A-18
; Sequence 18, Application US/10363332A
; Publication No. US20040038375A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; APPLICANT: No. US20040038375A1regard-Madsen, Mads
; TITLE OF INVENTION: Method For Screening Highly Active Proteases and Inhibitors
; FILE REFERENCE: 10080.204-US
; CURRENT APPLICATION NUMBER: US/10/363,332A
; CURRENT FILING DATE: 2003-05-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 2267
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Construct B

US-10-363-332A-18

Query Match 85.4%; Score 158; DB 12; Length 2267;
Best Local Similarity 97.0%; Pred. No. 8.5e-34;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 291

QY 134 GACATTATTTTACTGATATGATATAATAATTTGTATAAGAAAATG 179
DB 292 GTTATTTTACTGATATGATATAATAATTTGTATAAGAAAATG 337

RESULT 11
US-10-344-231-20
; Sequence 20, Application US/10344231
; Publication No. US20040038845A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; APPLICANT: Roggen, Erwin Ludo
; TITLE OF INVENTION: Improved production of proteases with inhibitors
; FILE REFERENCE: 10064.204-US
; CURRENT APPLICATION NUMBER: US/10/344,231
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 2588
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc feature
; LOCATION: (2070)..(2070)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2077)..(2077)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2084)..(2084)
; OTHER INFORMATION: n is a, c, g, or t
US-10-344-231-20

Query Match 85.4%; Score 158; DB 12; Length 2588;
Best Local Similarity 97.0%; Pred. No. 9e-34; 5; Indels 0; Gaps 0;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGGTT 291

QY 134 GACATTATTTTACTGATATGATATAATAATTTGTATAAGAAAATG 179
DB 292 GTTATTTTACTGATATGATATAATAATTTGTATAAGAAAATG 337

RESULT 12
US-10-363-332A-20
; Sequence 20, Application US/10363332A
; Publication No. US20040038375A1

GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; APPLICANT: No. US20040039375Alregaard-Madsen, Mads
; TITLE OF INVENTION: Method For Screening Highly Active Proteases and Inhibitors
; FILE REFERENCE: 10080.204-US
; CURRENT APPLICATION NUMBER: US/10/363.332A
; CURRENT FILING DATE: 2003-05-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 2588
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Construct D
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2070)..(2070)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2077)..(2077)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2084)..(2084)
; OTHER INFORMATION: n is a, c, g, or t
US-10-363-332A-20

Query Match 85.4%; Score 158; DB 12; Length 2588;
Best Local Similarity 97.0%; Pred. No. 9e-34; 5; Indels 0; Gaps 0;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 14 TCCAATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGTCTGTCATCAGACA 73
Db 172 TCGAATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGTCTGTCATCAGACA 231

Qy 74 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAATAAGTAAGGAGGAGGTT 133
Db 232 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAATAAGTAAGGAGGAGGTT 291

Qy 134 GACATATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 179
Db 292 GTATATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 337

RESULT 13
US-10-146-327-5
; Sequence 5, Application US/10146327
; Publication No. US2003004954A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Carsten
; APPLICANT: Jorgensen, Christel T.
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Svendsen, Allan
; APPLICANT: Kjaerulff, Soren
; TITLE OF INVENTION: Alpha-Amylase Variants
; FILE REFERENCE: 5886.200-US
; CURRENT APPLICATION NUMBER: US/10/146,327
; CURRENT FILING DATE: 2002-05-15
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: PA 1999 00437
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: 60/127,427
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2604
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
US-10-146-327-5

Query Match 85.0%; Score 157.2; DB 14; Length 2604;
Best Local Similarity 98.1%; Pred. No. 1.5e-33;
Matches 159; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 18 ATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 77
Db 593 ATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 652

Qy 78 ATTTTATGCTGCTCCAGACTGCTCCGCTGTGTAATAAGTAAGGAGGAGGTTGACA 137
Db 653 ATTTTATGCTGCTCCAGACTGCTCCGCTGTGTAATAAGTAAGGAGGAGGTTGTTA 712

Qy 138 TTATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 179
Db 713 TTATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 754

RESULT 14
US-09-769-864-58
; Sequence 58, Application US/09769864
; Patent No. US20010039253A1
; GENERAL INFORMATION:
; APPLICANT: Borchert, Torben V.
; APPLICANT: Svendsen, Allan
; APPLICANT: Andersen, Carsten
; APPLICANT: Nielsen, Bjarne
; APPLICANT: Nissen, Torben L.
; APPLICANT: Kjaerulff, Soren
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 5368.200-US
; CURRENT APPLICATION NUMBER: US/09/769,864
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/183,412
; PRIOR FILING DATE: 1998-10-30
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (343)...(1794)
US-09-769-864-58

Query Match 83.2%; Score 154; DB 9; Length 2084;
Best Local Similarity 96.9%; Pred. No. 1.1e-32;
Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 18 ATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 77
Db 73 ATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 132

Qy 78 ATTTTATGCTGCTCCAGACTGCTCCGCTGTGTAATAAGTAAGGAGGAGGTTGACA 137
Db 133 ATTTTATGCTGCTCCAGACTGCTCCGCTGTGTAATAAGTAAGGAGGAGGTTGTTA 192

Qy 138 TTATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 179
Db 193 TTATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 234

RESULT 15
US-09-854-346-9
; Sequence 9, Application US/09854346
; Patent No. US20020068352A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020068352A1ozymes A/S
; APPLICANT: Svendsen, Allan
; APPLICANT: Jorgensen, Christel Thea
; APPLICANT: Nielsen, Bjarne Ronfeldt
; TITLE OF INVENTION: Alpha-amylase variants with altered 1.6 activity

; FILE REFERENCE: 6140.200-US
; CURRENT APPLICATION NUMBER: US/09/854,346
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (343)..(1794)
; OTHER INFORMATION: BAN
US-09-854-346-9

Query Match 83.2%; Score 154; DB 9; Length 2084;
Best Local Similarity 96.9%; Pred.No.1.1e-32;
Matches 157; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 18 ATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTATCAGACAGGGT 77
Db 73 ATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTATCAGACAGGGT 132
Qy 78 ATTTTATGCTGCCAGACTGTCGGCTGTGTAATAAAATAGGAATAAAGGGGGTTGACA 137
Db 133 ATTTTATGCTGCCAGACTGTCGGCTGTGTAATAAAATAGGAATAAAGGGGGTTGTTA 192
Qy 138 TTATTTACTGATGATATATATAATTTGTATAGAAATG 179
Db 193 TTATTTACTGATGATATATAATTTGTATAGAAATG 234

Search completed: March 19, 2004, 05:16:24
Job time : 750.5 secs


```
XX SQ Sequence 185 BP; 62 A; 25 C; 42 G; 56 T; 0 U; 0 Other;
Query Match 100.0%; Score 185; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 3.1e-42;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCCCTGCAATCGATTGTTTGAGAAAAGAACCCATAAAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCCTGCAATCGATTGTTTGAGAAAAGAACCCATAAAAAATACCTTGT 60

QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGTCGCTGTGTAATAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGTCGCTGTGTAATAATAGGA 120

QY 121 ATAAAGGGGGTGCATATTTTACTGATATGTAATAATATTAATTTGTAAGAAATGG 180
DB 121 ATAAAGGGGGTGCATATTTTACTGATATGTAATAATATTAATTTGTAAGAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 2
AAF62647
ID AAF62647 standard; DNA; 185 BP.
XX AAF62647;
AC AAF62647;
XX 03-MAY-2001 (first entry)
XX Consensus amyQ promoter #2.
XX Pectin acetyltransferase; degrade; plant cell wall; ss.
XX Unidentified.
XX US6184028-B1.
XX 06-FEB-2001.
XX 26-AUG-1999; 99US-00384305.
XX 26-AUG-1999; 99US-00384305.
XX (NOVO ) NOVO NORDISK BIOTECH INC.
XX Thomas MD, Brown KM;
XX WPI; 2001-190946/19.
XX Novel isolated polypeptide having pectin acetyltransferase activity useful
XX for degrading pectic substances and in degradation or modification of
XX acetylated pectins and plant cell walls.
XX Example 9; Col 43-44; 35pp; English.
XX The present invention relates to Bacillus subtilis pectin acetyltransferase
XX protein. The invention is useful for degrading a pectic substance. It is
XX also useful for degrading soluble and insoluble pectins with varying
XX degrees of esterification, clarification etc. The protein may be used
XX alone or in combination with other enzymes for the degradation or
XX modification of acetylated pectins, degradation or modification of plant
XX cell walls
XX SQ Sequence 185 BP; 62 A; 25 C; 42 G; 56 T; 0 U; 0 Other;
Query Match 100.0%; Score 185; DB 4; Length 185;
Best Local Similarity 100.0%; Pred. No. 3.1e-42;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCCCTGCAATCGATTGTTTGAGAAAAGAACCCATAAAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCCTGCAATCGATTGTTTGAGAAAAGAACCCATAAAAAATACCTTGT 60

QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGTCGCTGTGTAATAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGTCGCTGTGTAATAATAGGA 120

QY 121 ATAAAGGGGGTGCATATTTTACTGATATGTAATAATATTAATTTGTAAGAAATGG 180
DB 121 ATAAAGGGGGTGCATATTTTACTGATATGTAATAATATTAATTTGTAAGAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 3
AAD09924
ID AAD09924 standard; DNA; 185 BP.
XX AAD09924;
AC AAD09924;
XX 12-SEP-2001 (first entry)
XX Bacillus amyloliquefaciens alpha-amylase (amyQ) consensus promoter #2.
XX Bacillus cell; tandem promoter; consensus promoter; mutant;
XX polypeptide production; amyL promoter; amyQ promoter; aprH promoter;
XX cryIIIA promoter; subtilisin Carlsberg gene promoter; alpha-amylase; ds.
XX Bacillus amyloliquefaciens.
XX Synthetic.
XX Key Location/Qualifiers
XX mutation replace(135, T)
XX /*tag= a
XX mutation replace(136, T)
XX /*tag= b
XX mutation replace(156, A)
XX /*tag= c
XX /note= "In column 26, a T to A mutation was said to occur
XX at position 116 but no such mutation was found in the
XX sequence shown in the sequence listing of the
XX specification"
XX US6255076-B1.
XX PN
XX 03-JUL-2001.
XX 26-FEB-1999; 99US-00258377.
XX 26-FEB-1998; 98US-00031442.
XX (NOVO ) NOVOZYMES BIOTECH INC.
XX Widner W, Sloma A, Thomas MD;
XX WPI; 2001-440518/47.
XX Producing a polypeptide in a Bacillus strain comprises cultivating a
XX Bacillus cell, which contains a nucleic acid construct comprising a
XX tandem promoter and consensus promoters.
XX Example 18; Col 43-44; 54pp; English.
XX The present invention relates to a method for producing a polypeptide,
XX comprises cultivating a Bacillus cell, which contains a nucleic acid
XX construct comprising a tandem promoter and consensus promoters. The
XX Bacillus is cultivated in a medium conducive for the production of the
XX polypeptide. Each promoter sequence of the tandem promoter is operably
XX linked to a nucleic acid sequence encoding the polypeptide. The tandem
XX promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIIA
XX promoter or subtilisin Carlsberg gene promoter. The consensus promoters
```

CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter or
CC subillisin Carlsberg gene promoter have the sequence TGGACA for the -35
CC region and TATAAT for the -10 region. The method further comprises
CC isolating the polypeptide from the cultivation medium. The method is
CC useful for producing a polypeptide in a Bacillus strain. The present
CC sequence is Bacillus amyloliquefaciens alpha-amylase (amyQ) consensus
CC mutated promoter, used in the exemplification of the invention

XX
SQ Sequence 185 BP; 62 A; 25 C; 42 G; 56 T; 0 U; 0 Other;
Query Match 100.0%; Score 185; DB 4; Length 185;
Best Local Similarity 100.0%; Pred. No. 3.1e-42; Indels 0; Gaps 0;
Matches 185; Conservative 0; Mismatches 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGT 60

QY 61 CTGTCTATCAGACAGGGTATTTTTATGCTGCTCCAGACTGTCGGCTGTGTAATAATAGGA 120
DB 61 CTGTCTATCAGACAGGGTATTTTTATGCTGCTCCAGACTGTCGGCTGTGTAATAATAGGA 120

QY 121 ATAAAGGGGGTGGACATTTATTTACTGATATGTAATAATAATTTGTATAAGAAATGG 180
DB 121 ATAAAGGGGGTGGACATTTATTTACTGATATGTAATAATAATTTGTATAAGAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 4
AAZ23323
ID AAZ23323 standard; DNA; 185 BP.
AC AAZ23323;
DT 06-DEC-1999 (first entry)
DE B. amyloliquefaciens amyQ promoter DNA #1.
XX Tandem promoter; consensus promoter; enzyme production; hormone; amyQ;
KW antibody; reporter; marker gene; cell regulation; alpha-amylase; ss.
XX Bacillus amyloliquefaciens.
XX WO943835-A2.
XX 02-SEP-1999.
XX 26-FEB-1999; 99WO-US004360.
XX 26-FEB-1998; 98US-00031442.
XX (NOVO) NOVO NORDISK BIOTECH INC.
XX Widner W, Sloma A, Thomas MD;
XX WPI; 1999-561370/47.
XX Production of polypeptide in Bacillus using specific promoters,
PT particularly for producing enzymes.
XX Claim 44; Page 90; 90pp; English.
XX This invention describes a novel method for the production of a
CC polypeptide in Bacillus using specific tandem or consensus promoters. The
CC method is used to produce homologous or particularly heterologous
CC proteins, particularly enzymes (specifically serine protease, maltogenic
CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
CC etc. The specified promoters provide increased expression of the sequence
CC which encodes the polypeptide of the invention. After incorporation of
CC the nucleic acid construct of the invention, any marker gene may be

CC deleted, resulting in a cell that is preferred for environmental and
CC regulatory regions. This sequence represents a Bacillus amyloliquefaciens
CC alpha-amylase amyQ promoter which is used to describe the method of the
CC invention

XX
SQ Sequence 185 BP; 63 A; 25 C; 42 G; 55 T; 0 U; 0 Other;
Query Match 99.1%; Score 183.4; DB 2; Length 185;
Best Local Similarity 99.5%; Pred. No. 8.8e-42; Indels 0; Gaps 0;
Matches 184; Conservative 0; Mismatches 1;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGT 60

QY 61 CTGTCTATCAGACAGGGTATTTTTATGCTGCTCCAGACTGTCGGCTGTGTAATAATAGGA 120
DB 61 CTGTCTATCAGACAGGGTATTTTTATGCTGCTCCAGACTGTCGGCTGTGTAATAATAGGA 120

QY 121 ATAAAGGGGGTGGACATTTATTTACTGATATGTAATAATAATTTGTATAAGAAATGG 180
DB 121 ATAAAGGGGGTGGACATTTATTTACTGATATGTAATAATAATTTGTATAAGAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 5
AAF62646
ID AAF62646 standard; DNA; 185 BP.
AC AAF62646;
DT 03-MAY-2001 (first entry)
DE Consensus amyQ promoter.
KW Pectin acetylase; degrade; plant cell wall; ss.
XX Unidentified.
XX US6184028-B1.
XX 06-FEB-2001.
XX 26-AUG-1999; 99US-00384305.
XX 26-AUG-1999; 99US-00384305.
XX (NOVO) NOVO NORDISK BIOTECH INC.
XX Thomas MD, Brown KM;
XX WPI; 2001-190946/19.
XX Novel isolated polypeptide having pectin acetylase activity useful
PT for degrading pectic substances and in degradation or modification of
PT acetylated pectins and plant cell walls.
XX Example 9; Col 43-44; 35pp; English.
XX The present invention relates to Bacillus subtilis pectin acetylase
CC protein. The invention is useful for degrading a pectic substance. It is
CC also useful for degrading soluble and insoluble pectins with varying
CC degrees of esterification, clarification etc. The protein may be used
CC alone or in combination with other enzymes for the degradation or
CC modification of acetylated pectins, degradation or modification of plant
CC cell walls
XX Sequence 185 BP; 63 A; 25 C; 42 G; 55 T; 0 U; 0 Other;
SQ Query Match 99.1%; Score 183.4; DB 4; Length 185;

CC which encodes the polypeptide of the invention. After incorporation of
CC the nucleic acid construct of the invention, any marker gene may be
CC deleted, resulting in a cell that is preferred for environmental and
CC regulatory regions. This sequence represents a *Bacillus thuringiensis*
CC alpha-amylase amyQ promoter which is used to describe the method of the
CC invention
XX
SQ Sequence 185 BP; 62 A; 24 C; 42 G; 57 T; 0 U; 0 Other;

Query Match 97.4%; Score 180.2; DB 2; Length 185;
Best Local Similarity 98.4%; Pred. No. 6.9e-41;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTCAATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTCAATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGT 60
QY 61 CTGTCATCAGACAGAGGTATTTTATGCTGTCACAGACTGTCGGCTGTGTAAGAAATAGGA 120
DB 61 CTGTCATCAGACAGAGGTATTTTATGCTGTCACAGACTGTCGGCTGTGTAAGAAATAGGA 120
QY 121 ATAAAGGGGGGTGACATTTTACTGATATGTAATAATTTGTATAGAAATAGG 180
DB 121 ATAAAGGGGGGTGTTTATTATTTTACTGATATGTAATAATTTGTATAGAAATAGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 8

AAF62645
ID AAF62645 standard; DNA; 185 BP.

XX AAF62645;

XX 03-MAY-2001 (first entry)

XX Wild type amyQ promoter.

XX Pectin acetyltransferase; degrade; plant cell wall; ss.

XX Unidentified.

XX US6184028-B1.

XX 06-FEB-2001.

XX 26-AUG-1999; 99US-00384305.

XX 26-AUG-1999; 99US-00384305.

XX (NOVO) NOVO NORDISK BIOTECH INC.

XX Thomas MD, Brown KM;

XX WPI; 2001-190946/19.

XX Novel isolated polypeptide having pectin acetyltransferase activity useful
XX for degrading pectic substances and in degradation or modification of
XX acetylated pectins and plant cell walls.

XX Example 9; Fig 9; 35pp; English.

XX The present invention relates to *Bacillus subtilis* pectin acetyltransferase
XX protein. The invention is useful for degrading a pectic substance. It is
XX also useful for degrading soluble and insoluble pectins with varying
XX degrees of esterification, clarification etc. The protein may be used
XX alone or in combination with other enzymes for the degradation or
XX modification of acetylated pectins, degradation or modification of plant
XX cell walls

XX Sequence 185 BP; 62 A; 24 C; 42 G; 57 T; 0 U; 0 Other;

Query Match 97.4%; Score 180.2; DB 4; Length 185;
Best Local Similarity 98.4%; Pred. No. 6.9e-41;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTCAATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGT 60

DB 1 GGCCTTAAGGGCGCTCAATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGT 60

QY 61 CTGTCATCAGACAGAGGTATTTTATGCTGTCACAGACTGTCGGCTGTGTAAGAAATAGGA 120

DB 61 CTGTCATCAGACAGAGGTATTTTATGCTGTCACAGACTGTCGGCTGTGTAAGAAATAGGA 120

QY 121 ATAAAGGGGGGTGACATTTTACTGATATGTAATAATTTGTATAGAAATAGG 180

DB 121 ATAAAGGGGGGTGTTTATTATTTTACTGATATGTAATAATTTGTATAGAAATAGG 180

QY 181 AGCTC 185

DB 181 AGCTC 185

RESULT 9

AAD09922
ID AAD09922 standard; DNA; 185 BP.

XX AAD09922;

XX 12-SEP-2001 (first entry)

XX *Bacillus amyloliquefaciens* alpha-amylase (amyQ) promoter.

XX *Bacillus* cell; tandem promoter; consensus promoter;

XX polypeptide production; amyL promoter; amyQ promoter; aprH promoter;

XX cryIIIA promoter; subtilisin Carlsberg gene promoter; alpha-amylase; ds.

XX *Bacillus amyloliquefaciens*.

XX US6255076-B1.

XX 03-JUL-2001.

XX 26-FEB-1999; 99US-00258377.

XX 26-FEB-1998; 98US-00031442.

XX (NOVO) NOVOZYMES BIOTECH INC.

XX Widner W, Sloma A, Thomas MD;

XX WPI; 2001-440518/47.

XX Producing a polypeptide in a *Bacillus* strain comprises cultivating a
XX *Bacillus* cell, which contains a nucleic acid construct comprising a
XX tandem promoter and consensus promoters.

XX Example 18; Col 43-44; 54pp; English.

XX The present invention relates to a method for producing a polypeptide,
XX comprises cultivating a *Bacillus* cell, which contains a nucleic acid
XX construct comprising a tandem promoter and consensus promoters. The
XX *Bacillus* is cultivated in a medium conducive for the production of the
XX polypeptide. Each promoter sequence of the tandem promoter is operably
XX linked to a nucleic acid sequence encoding the polypeptide. The tandem
XX promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIIA
XX promoter or subtilisin Carlsberg gene promoter. The consensus promoters
XX of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter or
XX subtilisin Carlsberg gene promoter have the sequence TTGACA for the -35
XX region and TATAAT for the -10 region. The method further comprises
XX isolating the polypeptide from the cultivation medium. The method is
XX useful for producing a polypeptide in a *Bacillus* strain. The present
XX sequence is *Bacillus amyloliquefaciens* alpha-amylase (amyQ) promoter,
XX used in the exemplification of the invention. Note: This sequence is

CC stated as being the same as that shown as SEQ ID NO 25 (AAD09931) in
CC figure 21 of the specification. However the sequences differ by at
CC position 116
XX
SQ Sequence 185 BP; 62 A; 24 C; 42 G; 57 T; 0 U; 0 Other;
Query Match 97.4%; Score 180.2; DB 4; Length 185;
Best Local Similarity 98.4%; Pred. No. 6.9e-41; Indels 0; Gaps 0;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGT 60
Db 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGAGGTAATTTTATGCTGCCAGACTGCCCTGTGTAATAATAGGA 120
Db 61 CTGTCATCAGACAGAGGTAATTTTATGCTGCCAGACTGCCCTGTGTAATAATAGGA 120
QY 121 ATAAAGGGGGTGTGATATTTTACTGATATGTAATAATAATTTCTATAAGAAATGG 180
Db 121 ATAAAGGGGGTGTGATATTTTACTGATATGTAATAATAATTTCTATAAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185
RESULT 10
AAZ23325
ID AAZ23325 standard; DNA; 185 BP.
XX
AC AAZ23325;
XX
DT 06-DEC-1999 (first entry)
XX
DE Bacillus sp. amyQ promoter consensus DNA.
XX
KW Tandem promoter; consensus promoter; enzyme production; hormones; amyQ;
KW antibody; reporter; marker gene; cell regulation; alpha-amylase; ss.
XX
OS Bacillus sp.
XX
PN WO9943835-A2.
XX
PD 02-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US004360.
XX
PR 26-FEB-1998; 98US-00031442.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
XX
PI Widner W, Sloma A, Thomas MD;
XX
DR WPI; 1999-561370/47.
XX
PT Production of polypeptide in Bacillus using specific promoters,
XX particularly for producing enzymes.
XX
PS Disclosure; Fig 21; 90pp; English.
XX
CC This invention describes a novel method for the production of a
CC polypeptide in Bacillus using specific tandem or consensus promoters. The
CC method is used to produce homologous or particularly heterologous
CC proteins, particularly enzymes (specifically serine protease, maltogenic
CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
CC etc. The specified promoters provide increased expression of the sequence
CC which encodes the polypeptide of the invention. After incorporation of
CC the nucleic acid construct of the invention, any marker gene may be
CC deleted, resulting in a cell that is preferred for environmental and
CC regulatory regions. This sequence represents the consensus sequence of a
CC Bacillus sp. alpha-amylase amyQ promoter which is used to describe the
CC method of the invention

XX
SQ Sequence 185 BP; 63 A; 24 C; 42 G; 56 T; 0 U; 0 Other;
Query Match 96.5%; Score 178.6; DB 2; Length 185;
Best Local Similarity 97.8%; Pred. No. 1.9e-40;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGT 60
Db 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGAGGTAATTTTATGCTGCCAGACTGCCCTGTGTAATAATAGGA 120
Db 61 CTGTCATCAGACAGAGGTAATTTTATGCTGCCAGACTGCCCTGTGTAATAATAGGA 120
QY 121 ATAAAGGGGGTGTGATATTTTACTGATATGTAATAATAATTTCTATAAGAAATGG 180
Db 121 ATAAAGGGGGTGTGATATTTTACTGATATGTAATAATAATTTCTATAAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185
RESULT 11
AAD09931
ID AAD09931 standard; DNA; 185 BP.
XX
AC AAD09931;
XX
DT 12-SEP-2001 (first entry)
XX
DE Bacillus amyloliquefaciens amyQ promoter, alternative version.
XX
KW Bacillus cell; tandem promoter; consensus promoter;
KW polypeptide production; amyL promoter; amyQ promoter; apxH promoter;
KW cryIIIA promoter; subtilisin Carlsberg gene promoter; alpha-amylase; ds.
XX
OS Bacillus amyloliquefaciens.
XX
FH Key Location/Qualifiers
FT misc_feature 116
FT /tag= a
FT /note= "Base T is found at this location in the sequence
FT shown in the sequence listing of the specification
FT (AAD09922)"
XX
US6255076-B1.
XX
PN
XX
PD 03-JUL-2001.
XX
PF 26-FEB-1999; 99US-00258377.
XX
PR 26-FEB-1998; 98US-00031442.
XX
PA (NOVO) NOVOZYMES BIOTECH INC.
XX
PI Widner W, Sloma A, Thomas MD;
XX
DR WPI; 2001-440518/47.
XX
PT Producing a polypeptide in a Bacillus strain comprises cultivating a
PT Bacillus cell, which contains a nucleic acid construct comprising a
PT tandem promoter and consensus promoters.
XX
PS Example 18; Fig 21; 54pp; English.
XX
CC The present invention relates to a method for producing a polypeptide,
CC comprises cultivating a Bacillus cell, which contains a nucleic acid
CC construct comprising a tandem promoter and consensus promoters. The
CC Bacillus is cultivated in a medium conducive for the production of the
CC polypeptide. Each promoter sequence of the tandem promoter is operably
CC linked to a nucleic acid sequence encoding the polypeptide. The tandem

XX Screening protease/protease inhibitor gene library for gene encoding the
PT same, by constructing host cell having the gene, cultivating cell,
PT dissociating complex formed between them, selecting protein and isolating

XX	09-JUN
PF	
XX	
PR	21-AUG
PR	24-AUG
XX	
PA	(NOV)
XX	

XX	Synthetic.	XX
OS	Bacillus sp.	XX
OS	Hordeum sp.	XX
OS		XX
XX	WO200216619-A1.	XX
PN		XX
XX	28-FEB-2002.	XX
PD		XX
XX	09-JUL-2001; 2001WO-DK000479.	XX
XX		XX
PR	21-AUG-2000; 2000DK-00001233.	XX
PR	24-AUG-2000; 2000US-0227501P.	XX
XX		XX
PA	(NOVO) NOVOZYMES AS.	XX
XX		XX

PI Pedersen PE;
XX WPI; 2002-242031/29.
XX Producing protease-inhibitor compositions with reduced allergenicity,
PT stability and activity for use in detergent compositions.
XX
PS Example 1; Page 45-46; 52pp; English.
XX
CC The invention provides a method for producing a protease-inhibitor
CC complex with reduced allergenicity for use in detergents. The method
CC involves (a) constructing a fusion polynucleotide sequence in frame, the
CC sequence comprising a first gene encoding a protease (preferably a
CC subtilase S11 or S12) and a second gene encoding a protease inhibitor;
CC (b) introducing the sequence into a host cell; and (c) cultivating the
CC host cell (the host cell expresses the sequence and produces a non-
CC covalently linked complex of the protease and the inhibitor). The
CC protease-inhibitor complex produced by the method may be used as an
CC nucleotide sequence of a DNA construct comprising the savinase-CI-2A
CC fusion sequence
XX
SQ Sequence 2166 BP; 610 A; 442 C; 582 G; 532 T; 0 U; 0 Other;
Query Match 85.4%; Score 158; DB 6; Length 2166;
Best Local Similarity 97.0%; Pred. No. 2e-34;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATAAAATACCTTGTCTCTCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAGAGAGACCATAAAATACCTTGTCTCTCATCAGACA 231
QY 74 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAATAAATAGGAATAAGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAATAAATAGGAATAAGGGGGTT 291
QY 134 GACATTATTTACTGATATGTAATAATAATTTGTATAAGAAATG 179
DB 292 GTTATTATTTACTGATATGTAATAATAATTTGTATAAGAAATG 337
RESULT 14
AAD34434
ID AAD34434 standard; DNA; 2267 BP.
XX AAD34434;
AC AAD34434;
DT 29-AUG-2003 (revised)
DT 16-JUL-2002 (first entry)
DE Subtilisin 309-CI-2A transcriptional product encoding construct B DNA.
XX
KW Protease; screening; barley; chymotrypsin inhibitor; CI-2A; laundry;
KW detergent industry; dishwash; fabric softener; subtilisin 309; savinase;
KW chimeric; ds.
XX
OS Hordeum sp.
OS Bacillus lentus.
OS Chimeric.
XX
XX WO200216598-A1.
XX
XX 07-MAR-2002.
XX
XX 17-JUL-2001; 2001WO-DK000503.
XX
XX 29-AUG-2000; 2000DK-00001273.
PR 05-SEP-2000; 2000US-0230017P.
XX
XX (NOVO) NOVOZYMES AS.
XX
XX Pedersen PE, Norregaard-Madsen M;
PI WPI; 2002-329776/36.
DR

XX Screening protease/protease inhibitor gene library for gene encoding the
PT same, by constructing host cell having the gene, cultivating cell,
PT dissociating complex formed between them, selecting protein and isolating
PT gene.
XX
XX Example 1; Page 48; 54pp; English.
XX
CC The invention relates to a method for screening a protease/protease
CC inhibitor gene library for a gene encoding a protease of interest. The
CC method comprises introducing a first gene of the protease gene library
CC and a second gene encoding a protease inhibitor into a host cell;
CC cultivating the host cell, wherein the cell expresses the first and the
CC second genes to produce a complex of a protease and the inhibitor;
CC dissociating the inhibitor from the complex and selecting the protease of
CC interest and isolating the encoding gene. A complex may be used for a
CC number of industrial applications, in particular within the detergent
CC industry, thus, relates to a cleaning or detergent composition,
CC preferably a laundry or dishwash composition comprising the complex,
CC where hand or machine laundry detergent composition including a laundry
CC additive composition suitable for pre-treatment of stained fabrics and a
CC rinse added composition, or be formulated as a detergent composition for
CC use in general household hard surface cleaning operations or be
CC formulated for hand or machine dish washing operations. The present
CC sequence is a construct encoding a transcriptional fusion product
CC containing C-terminal region of subtilisin 309 (savinase) and barley
CC chymotrypsin inhibitor CI-2A protein used in the invention. (Updated on
CC 29-AUG-2003 to standardise CS field)
XX
SQ Sequence 2267 BP; 640 A; 462 C; 601 G; 564 T; 0 U; 0 Other;
Query Match 85.4%; Score 158; DB 6; Length 2267;
Best Local Similarity 97.0%; Pred. No. 2e-34;
Matches 161; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 14 TGCATCGATTGTTTGAGAAAGAGAGACCATAAAATACCTTGTCTCTCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAGAGAGACCATAAAATACCTTGTCTCTCATCAGACA 231
QY 74 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAATAAATAGGAATAAGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGCCGCTGTGTAATAAATAGGAATAAGGGGGTT 291
QY 134 GACATTATTTACTGATATGTAATAATAATTTGTATAAGAAATG 179
DB 292 GTTATTATTTACTGATATGTAATAATAATTTGTATAAGAAATG 337
RESULT 15
ABL40499
ID ABL40499 standard; DNA; 2267 BP.
XX ABL40499;
AC ABL40499;
XX
XX 17-JUN-2002 (first entry)
XX
XX DNA construct B comprising Savinase-CI-2A fusion sequence.
DE
DE Protease; allergenicity; detergent; subtilase; S11; S12; additive;
KW protease inhibitor; barley; chymotrypsin; CI-2A; fusion protein;
KW subtilisin; savinase; apr gene; ds.
XX
XX Synthetic.
OS Bacillus sp.
OS Hordeum sp.
XX
XX WO200216619-A1.
XX
XX 28-FEB-2002.
XX
XX 09-JUL-2001; 2001WO-DK000479.
PF
XX 21-AUG-2000; 2000DK-00001233.
PR

```

PR 24-AUG-2000; 2000US-0227501P.
XX
XX PA (NOVO ) NOVOZYMES AS.
XX
XX PI Pedersen PE;
XX
XX DR WPI; 2002-242031/29.
XX
XX PT Producing protease-inhibitor compositions with reduced allergenicity,
XX PT stability and activity for use in detergent compositions.
XX
XX PS Example 1; Page 46; 52pp; English.
XX
XX CC The invention provides a method for producing a protease-inhibitor
XX CC complex with reduced allergenicity for use in detergents. The method
XX CC involves (a) constructing a fusion polynucleotide sequence in frame, the
XX CC sequence comprising a first gene encoding a protease (preferably a
XX CC subtilase S11 or S12) and a second gene encoding a protease inhibitor;
XX CC (b) introducing the sequence into a host cell; and (c) cultivating the
XX CC host cell (the host cell expresses the sequence and produces a non-
XX CC covalently linked complex of the protease and the inhibitor). The
XX CC protease-inhibitor complex produced by the method may be used as an
XX CC nucleotide sequence of a DNA construct comprising the savinase-CI-2A
XX CC fusion sequence
XX
XX SQ Sequence 2267 BP; 640 A; 462 C; 601 G; 564 T; 0 U; 0 Other;

Query Match      85.4%; Score 158; DB 6; Length 2267;
Best Local Similarity 97.0%; Fred. No. 2e-34; Mismatches 5; Indels 0; Gaps 0;
Matches 161; Conservative 0;

QY 14 TGCATCGATTGTTTCAGAAAAGAGAGACCATAAATAACCTTGTCTGTCATCAGACA 73
   |||||||
Db 172 TGCATCGATTGTTTCAGAAAAGAGAGACCATAAATAACCTTGTCTGTCATCAGACA 231

QY 74 GGGTATTTTATGCTGTCCGACTGTCGCTGTGTAAATAAGGAGGAGGTT 133
   |||||||
Db 232 GGGTATTTTATGCTGTCCGACTGTCGCTGTGTAAATAAGGAGGAGGTT 291

QY 134 GACATTATTCTACTGATATGATATAATATTGTATAAGAAAATG 179
   |||||||
Db 292 GTTATTATTCTACTGATATGATATAATATTGTATAAGAAAATG 337

```

Search completed: March 19, 2004, 01:38:17
Job time : 215 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 19, 2004, 00:23:20 ; Search time 1078.5 Seconds
(without alignments)
7434.827 Million cell updates/sec

Title: US-09-834-271A-27
Perfect score: 185
Sequence: 1 gsccttaaggcctgaatc.....tgtataagaatggagctc 185

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sv.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_man.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	185	100.0	185	6	AR160503	AR160503 Sequence
2	185	100.0	185	6	AX088921	AX088921 Sequence
3	185	100.0	185	6	BD136799	BD136799 Process f
4	183.4	99.1	185	6	AR160502	AR160502 Sequence
5	183.4	99.1	185	6	AX088920	AX088920 Sequence
6	183.4	99.1	185	6	BD136798	BD136798 Process f
7	180.2	97.4	185	6	AR160501	AR160501 Sequence
8	180.2	97.4	185	6	AX088919	AX088919 Sequence
9	180.2	97.4	185	6	BD136797	BD136797 Process f
10	157.2	85.0	249	6	BD174625	BD174625 Modified
11	157.2	85.0	2604	6	AR168304	AR168304 Sequence
12	157.2	85.0	2604	6	BD249244	BD249244 Alpha-amy
13	157.2	85.0	2604	6	AR215288	AR215288 Sequence
14	157.2	85.0	2604	6	AX036892	AX036892 Sequence
15	157.2	85.0	2604	6	BD137366	BD137366 Enzymatic
16	154	83.2	162	6	I40596	I40596 Sequence 24
17	154	83.2	162	6	I40867	I40867 Sequence 24
18	154	83.2	162	6	I40895	I40895 Sequence 24
19	154	83.2	162	6	I56837	I56837 Sequence 24
20	154	83.2	537	1	BAAMYL	V00092 Bacillus am
21	154	83.2	576	6	A00155	A00155 Nucleotide
22	154	83.2	576	6	A00156	A00156 Nucleotide
23	154	83.2	576	6	A00157	A00157 Nucleotide
24	154	83.2	576	6	A00158	A00158 Nucleotide
25	154	83.2	576	6	A00604	A00604 B.amyloliqu
26	154	83.2	576	6	A00605	A00605 B.amyloliqu
27	154	83.2	2084	1	BACRAM	J01542 Bacillus am
28	154	83.2	2084	6	AR068285	AR068285 Sequence
29	154	83.2	2084	6	AR037275	AR037275 Sequence
30	154	83.2	2084	6	AR052145	AR052145 Sequence
31	154	83.2	2084	6	AR087551	AR087551 Sequence
32	154	83.2	2084	6	AR129916	AR129916 Sequence
33	154	83.2	2084	6	AR137905	AR137905 Sequence
34	154	83.2	2084	6	AR143262	AR143262 Sequence
35	154	83.2	2084	6	BD230131	BD230131 Alpha-amy
36	154	83.2	2084	6	AR224261	AR224261 Sequence
37	154	83.2	2084	6	AR285306	AR285306 Sequence
38	154	83.2	2084	6	AR428861	AR428861 Sequence
39	154	83.2	2084	6	AX244195	AX244195 Sequence
40	154	83.2	2084	6	AX305020	AX305020 Sequence
41	154	83.2	2084	6	AX339268	AX339268 Sequence
42	154	83.2	2084	6	AX370717	AX370717 Sequence
43	154	83.2	2084	6	AX601810	AX601810 Sequence
44	154	83.2	2084	6	BD077082	BD077082 Alpha-amy
45	154	83.2	2084	6	BD085805	BD085805 Alpha-amy

ALIGNMENTS

RESULT 1
AR160503
LOCUS
DEFINITION Sequence 27 from patent US 6255076.
ACCESSION AR160503
VERSION AR160503.1 GI:16224480
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 185)
AUTHORS Widner,W., Sloma,A. and Thomas,M.D.
TITLE Methods for producing a polypeptide in a Bacillus cell
JOURNAL Patent: US 6255076-A 27 03-JUL-2001;
FEATURES Location/Qualifiers

AR160503 185 bp DNA linear PAT 17-OCT-2001

```
source 1..185
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 185; DB 6; Length 185;
Best Local Similarity 100.0%; Pred. No. 9.8e-36;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCTCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAGGA 120
QY 121 ATAAAGGGGGTGCATTATTTTACTGATATGATATATAATTTGTATAAGAAATGG 180
DB 121 ATAAAGGGGGTGCATTATTTTACTGATATGATATATAATTTGTATAAGAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 2
AX088921 185 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 22 from Patent WO0114534.
ACCESSION AX088921
VERSION AX088921.1 GI:13397680
KEYWORDS
SOURCE
ORGANISM Bacillus subtilis
REFERENCE
AUTHORS Thomas,M.D. and Brown,K.M.
TITLE Polypeptides having pectin acetylesterase activity and nucleic acids encoding same
JOURNAL Patent: WO 0114534-A 22 01-MAR-2001;
FEATURES
source Novozymes Biotech, Inc. (US)
1..185
/organism="Bacillus subtilis"
/mol_type="unassigned DNA"
/db_xref="taxon:1423"

ORIGIN
Query Match 100.0%; Score 185; DB 6; Length 185;
Best Local Similarity 100.0%; Pred. No. 9.8e-36;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCTCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAGGA 120
QY 121 ATAAAGGGGGTGCATTATTTTACTGATATGATATATAATTTGTATAAGAAATGG 180
DB 121 ATAAAGGGGGTGCATTATTTTACTGATATGATATATAATTTGTATAAGAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 3
BD136799 185 bp DNA linear PAT 18-SEP-2002
LOCUS
```

```
DEFINITION Process for producing polypeptide in bacillus cells.
ACCESSION BD136799
VERSION BD136799.1 GI:23231744
KEYWORDS JP 2002504379-A/27.
SOURCE unidentified
ORGANISM unidentified
COMMENT 1 (bases 1 to 185)
REFERENCE Widner,W., Sloma,A. and Thomas,M.D.
AUTHORS Process for producing polypeptide in bacillus cells
TITLE Patent: JP 2002504379-A 27 12-FEB-2002;
JOURNAL NOVO NORDISK BIOTECH INC
COMMENT OS Bacillus
PN JP 2002504379-A/27
PD 12-FEB-2002
PF 26-FEB-1999 JP 2000533574
PR 26-FEB-1998 US 09/031442
PI WILLIAM WIDNER,ALAN SLOMA,MICHAEL D THOMAS
PC C12N15/09,C12N1/21,C12N9/54//C07K14/325,(C12N1/21,C12R1:07),
PC C12N15/00
CC Process for producing polypeptide in bacillus cells FH Key
FT source 1..185
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

FEATURES
source
1..185
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 185; DB 6; Length 185;
Best Local Similarity 100.0%; Pred. No. 9.8e-36;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCTCAATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAGGA 120
QY 121 ATAAAGGGGGTGCATTATTTTACTGATATGATATATAATTTGTATAAGAAATGG 180
DB 121 ATAAAGGGGGTGCATTATTTTACTGATATGATATATAATTTGTATAAGAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 4
AR160502 185 bp DNA linear PAT 17-OCT-2001
LOCUS Sequence 26 from patent US 6255076.
ACCESSION AR160502
VERSION AR160502.1 GI:16224477
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
COMMENT 1 (bases 1 to 185)
REFERENCE Widner,W., Sloma,A. and Thomas,M.D.
AUTHORS Methods for producing a polypeptide in a Bacillus cell
TITLE Patent: US 6255076-A 26 03-JUL-2001;
JOURNAL Location/Qualifiers
FEATURES
source 1..185
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 99.1%; Score 183.4; DB 6; Length 185;
```

```
Best Local Similarity 99.1%; Pred. No. 2.4e-35; Mismatches 1; Indels 0; Gaps 0;
Matches 184; Conservative 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCACTAAATACCTTGT 60
Db 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCACTAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAATAGGA 120
Db 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAATAGGA 120
QY 121 ATAAAGGGGGTGGACATATTTTACTGATATGTAATATATATATTTGTAAGAAATGG 180
Db 121 ATAAAGGGGGTGGACATATTTTACTGATATGTAATATATATATTTGTAAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 5
AX088920
LOCUS AX088920 185 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 21 from Patent WO0114534.
ACCESSION AX088920
VERSION AX088920.1 GI:13397679
KEYWORDS
SOURCE
ORGANISM
Bacillus subtilis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
1 Thomas,M.D. and Brown,K.M.
Polypeptides having pectin acetyltransferase activity and nucleic
acids encoding same
JOURNAL Patent: WO 0114534-A 21 01-MAR-2001;
Novozymes Biotech, Inc. (US)
FEATURES
source
1. .185
/organism="Bacillus subtilis"
/mol_type="unassigned DNA"
/db_xref="taxon:1423"

ORIGIN
Query Match 99.1%; Score 183.4; DB 6; Length 185;
Best Local Similarity 99.5%; Pred. No. 2.4e-35;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCACTAAATACCTTGT 60
Db 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCACTAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAATAGGA 120
Db 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAATAGGA 120
QY 121 ATAAAGGGGGTGGACATATTTTACTGATATGTAATATATATATTTGTAAGAAATGG 180
Db 121 ATAAAGGGGGTGGACATATTTTACTGATATGTAATATATATATTTGTAAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 6
BD136798
LOCUS BD136798 185 bp DNA linear PAT 18-SEP-2002
DEFINITION Process for producing polypeptide in bacillus cells.
ACCESSION BD136798
VERSION BD136798.1 GI:23231743
KEYWORDS
SOURCE JP 2002504379-A/26.
unidentified
ORGANISM
```

```
unclassified.
1 (bases 1 to 185)
Widner,W., Sloma,A. and Thomas,M.D.
Process for producing polypeptide in bacillus cells
Patent: JP 2002504379-A 26 12-FEB-2002;
NOVO NORDISK BIOTECH INC
OS Bacillus
PN JP 2002504379-A/26
PD 12-FEB-2002
PF 26-FEB-1999 JP 2000533574
PR 26-FEB-1998 US 09/031442
PI WILLIAM WIDNER,ALAN SLOMA,MICHAEL D THOMAS
PC C12N15/09,C12N1/21,C12N9/54//C07K14/325,(C12N1/21,C12N1:07),
C12N15/00
CC Process for producing polypeptide in bacillus cells FH Key
FT source
1. .185
Location/Qualifiers
/organism="Bacillus".
FEATURES
source
1. .185
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32844"

ORIGIN
Query Match 99.1%; Score 183.4; DB 6; Length 185;
Best Local Similarity 99.5%; Pred. No. 2.4e-35;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCACTAAATACCTTGT 60
Db 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCACTAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAATAGGA 120
Db 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAATAGGA 120
QY 121 ATAAAGGGGGTGGACATATTTTACTGATATGTAATATATATATTTGTAAGAAATGG 180
Db 121 ATAAAGGGGGTGGACATATTTTACTGATATGTAATATATATATTTGTAAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 7
AR160501
LOCUS AR160501 185 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 25 from patent US 6255076.
ACCESSION AR160501
VERSION AR160501.1 GI:16224474
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
1 (bases 1 to 185)
Widner,W., Sloma,A. and Thomas,M.D.
Methods for producing a polypeptide in a Bacillus cell
Patent: US 6255076-A 25 03-JUL-2001;
FEATURES
source
1. .185
Location/Qualifiers
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 97.4%; Score 180.2; DB 6; Length 185;
Best Local Similarity 98.4%; Pred. No. 1.5e-34;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCACTAAATACCTTGT 60
Db 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAACCACTAAATACCTTGT 60
```

QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAATAGGA 120
QY 121 ATAAAGGGGGTTCACATTATTTTACTGATATGTATATATATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTCACATTATTTTACTGATATGTATATATATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 8
LOCUS AX088919 185 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 20 from Patent WO0114534.
ACCESSION AX088919
VERSION AX088919.1 GI:13397678
KEYWORDS Bacillus subtilis
SOURCE Bacillus subtilis
ORGANISM Bacillus subtilis
REFERENCE 1
AUTHORS Thomas, M.D. and Brown, K.M.
TITLE Polypeptides having pectin acetylesterase activity and nucleic acids encoding same
JOURNAL Patent: WO 0114534-A 20 01-MAR-2001;
Novozymes Biotech, Inc. (US)
FEATURES
source 1.185
/organism="Bacillus subtilis"
/mol_type="unassigned DNA"
/db_xref="taxon:1423"

ORIGIN
Query Match 97.4%; Score 180.2; DB 6; Length 185;
Best Local Similarity 98.4%; Pred. No. 1.5e-34;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCCCTGCAATCGATTGTTGAGAAAGAGAGACCACATAAAAATACCTTGT 60
DB 1 GGCCTTAAGGCCCTGCAATCGATTGTTGAGAAAGAGAGACCACATAAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAATAGGA 120
QY 121 ATAAAGGGGGTTCACATTATTTTACTGATATGTATATATATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTCACATTATTTTACTGATATGTATATATATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 9
LOCUS BD136797 185 bp DNA linear PAT 18-SEP-2002
DEFINITION Process for producing polypeptide in bacillus cells.
ACCESSION BD136797
VERSION BD136797.1 GI:23231742
KEYWORDS JP 2002504379-A/25.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 185)
AUTHORS Widner, W., Sloma, A. and Thomas, M.D.
TITLE Process for producing polypeptide in bacillus cells
JOURNAL Patent: JP 2002504379-A 25. 12-FEB-2002;
NOVO NORDISK BIOTECH INC

COMMENT OS Bacillus
FN JP 2002504379-A/25
PD 12-FEB-2002
PF 26-FEB-1999 JP 2000533574
PI WILLIAM WIDNER, ALAN SLOMA, MICHAEL D THOMAS
PC C12N15/09, C12N1/21, C12N9/54, C07K14/325, (C12N1/21, C12R1:07),
C12N15/00
CC Process for producing polypeptide in bacillus cells FH Key
FT source 1.185
FT Location/Qualifiers
1.185
Location/Qualifiers
/organism="Bacillus".
1.185
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 97.4%; Score 180.2; DB 6; Length 185;
Best Local Similarity 98.4%; Pred. No. 1.5e-34;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCCCTGCAATCGATTGTTGAGAAAGAGAGACCACATAAAAATACCTTGT 60
DB 1 GGCCTTAAGGCCCTGCAATCGATTGTTGAGAAAGAGAGACCACATAAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAATAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAATAGGA 120
QY 121 ATAAAGGGGGTTCACATTATTTTACTGATATGTATATATATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTCACATTATTTTACTGATATGTATATATATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 10
LOCUS BD174625 249 bp DNA linear PAT 18-MAR-2003
DEFINITION Modified promoter.
ACCESSION BD174625
VERSION BD174625.1 GI:29120315
KEYWORDS JP 200272466-A/14.
SOURCE Bacillus amyloliquefaciens
ORGANISM Bacillus amyloliquefaciens
REFERENCE 1 (bases 1 to 249)
AUTHORS Bacteri; Firmicutes; Bacillales; Bacillaceae; Bacillus.
Mizubuchi, H., Fushimi, N. and Miyoshi, S.
TITLE Modified promoter
JOURNAL Patent: JP 200272466-A 14 24-SEP-2002;
SHOWA SANGYO CO LTD
COMMENT OS Bacillus amyloliquefaciens
FN JP 200272466-A/14
PD 24-SEP-2002
PF 15-MAR-2001 JP 2001074780
PI HIROYUKI MIZUBUCHI, NAOKA FUSHIMI, SHINSUKE MIYOSHI PC
C12N15/09, C12N1/21, C12P21/02, (C12N1/21, C12R1:07), (C12N1/21, C12R1:01),
PC (C12N1/21, C12R1:19), C12N15/00
CC Modified promoter
FH Key Location/Qualifiers
FT source 1.249
FT Location/Qualifiers
/organism="Bacillus amyloliquefaciens".
1.249
/organism="Bacillus amyloliquefaciens"
/mol_type="genomic DNA"
/db_xref="taxon:1390"

ORIGIN

PD	03-DEC-2002
PF	28-MAR-2000 JP 2006069551
PR	30-MAR-1999 DK PA 199900437
PI	CARSTEN ANDERSEN,CHRISTEL THEA JOERGENSEN,HENRIK BISGARD PI FRANTZEN,
PI	ALLAN SVENDSEN,SOEREN KJARRULFF
PC	C12N15/09,C11D3/386,C12N1/15,C12N1/19,C12N5/10,C12N9/
PC	28,C12P7/06,
PC	C12P19/14,C12S11/00//C12N9/28,C12R1:07),(C12N9/28,C12R1:10),
PC	(C12N9/28,C12R1:125),C12N5/00,C12N5/00
CC	Alpha-amylase variants
PH	Key Location/Qualifiers
FT	-10 signal (707)..(712)
FT	-35_signal (729)..(734)
FT	RBS (759)..(762)
FT	sig_peptide (770)..(862)
FT	mat_peptide (863)..(2314)
FT	terminator (2321)..(2376)
FT	CDS (863)..(2314).
FEATURES	Location/Qualifiers
source	1..2604
	/organism="Bacillus amyloliquefaciens"
	/mol_type="genomic DNA"
	/db_xref="taxon:1390"
ORIGIN	
Query Match	85.0%; Score 157.2; DB 6; Length 2604;
Best Local Similarity	98.1%; Pred.No.5.3e-29;
Matches 159; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
QY	18 ATCGATTGGTTTGAGAAAAGAAGAACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 77
Db	593 ATCGATTGGTTTGAGAAAAGAAGAACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 652
QY	78 ATTTTTATGCTGTCAGACTGCCGCTGTGTAAAAATAGGAATAAGGGGGTTTGACA 137
Db	653 ATTTTTATGCTGTCAGACTGCCGCTGTGTAAAAATAGGAATAAGGGGGTTTGTTA 712
QY	138 TTATTTTACTGATCATGTAATAATATTTGTATAAGAAAATG 179
Db	713 TTATTTTACTGATCATGTAATAATATTTGTATAAGAAAATG 754
RESULT 13	
AR215288	
LOCUS	AR215288 2604 bp DNA linear PAT 25-SEP-2002
DEFINITION	Sequence 5 from patent US 6410295.
ACCESSION	AR215288
VERSION	AR215288.1 GI:23313486
KEYWORDS	
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 2604)
TITLE	Andersen,C., Jorgensen,C.T., Bisgrd-Frantzen,H., Svendsen,A. and
JOURNAL	Kjaerulf,S.
FEATURES	Alpha-amylase variante
source	Patent: US 6410295-A 5 25-JUN-2002;
	Location/Qualifiers
	1..2604
	/organism="unknown"
	/mol_type="genomic DNA"
ORIGIN	
Query Match	85.0%; Score 157.2; DB 6; Length 2604;
Best Local Similarity	98.1%; Pred.No.5.3e-29;
Matches 159; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
QY	18 ATCGATTGGTTTGAGAAAAGAAGAACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 77
Db	593 ATCGATTGGTTTGAGAAAAGAAGAACCAATAAAATACCTTGTCTGTCATCAGACAGGTT 652
QY	78 ATTTTTATGCTGTCAGACTGCCGCTGTGTAAAAATAGGAATAAGGGGGTTTGACA 137

[illegible]

LOCUS	BD137366	2604 bp	DNA	linear	PAT 18-SEP-2000			
DEFINITION	Enzymatic preparation of glucose syrup from starch.							
ACCESSION	BD137366							
VERSION	BD137366.1	GI:23232311						
KEYWORDS	JP 2002505885-A/2.							
SOURCE	Bacillus amyloliquefaciens							
ORGANISM	Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.							
REFERENCE	1 (bases 1 to 2604)							
AUTHORS	Norman, B.E. and Hendriksen, H.V.							
TITLE	Enzymatic preparation of glucose syrup from starch							
JOURNAL	Patent: JP 2002505885-A 26-FEB-2002;							
COMMENT	NOVOZYMES AS							
	OS	Bacillus amyloliquefaciens						
	PN	JP 2002505885-A/2						
	PD	26-FEB-2002						
	PF	08-MAR-1999	JP 2000535766					
	PR	09-MAR-1998	DK 0321/98					
	PI	BARRIE EDMUND NORMAN, HANNE VANG HENDRIKSEN						
	PC	C12P19/14, C12N9/26, C12N15/09, C13K1/06, C12N15/00	CC	Enzymatic				
	preparation of glucose syrup from starch FH Key							
	Location/Qualifiers							
	FT	-10 signal	(707) . . (712)					
	FT	-35_signal	(729) . . (734)					
	FT	RBS	(759) . . (762)					
	FT	sig_peptide	(770) . . (862)					
	FT	mat_peptide	(863) . . (2314)					
	FT	terminator	(2321) . . (2376)					
	FT	CDS	(770) . . (2314)					
FEATURES	Location/Qualifiers							
source	1. 2604							
	/organism="Bacillus amyloliquefaciens"							
	/mol_type="genomic DNA"							
	/db_xref="taxon:1390"							
ORIGIN								
Query Match	85.0%; Score 157.2; DB 6; Length 2604;							
Best Local Similarity	98.1%; Pred. No. 5.3e-29;							
Matches 159;	Conservative	0;	Mismatches	3;	Indels 0; Gaps 0;			
Qy	18	ATCGATTGTTTGAGAAAAGAGAACCCATAAAAAATACCTTGTCTGTCTATCAGCAGGGT	77					
Db	593	ATCGATTGTTTGAGAAAAGAGAACCCATAAAAAATACCTTGTCTGTCTATCAGCAGGGT	652					
Qy	78	ATTTTATTCCTGCTCCAGACTGTCCTGTGTAAAAAATAGGAATAAAGGGGGTTTGACA	137					
Db	653	ATTTTATTCCTGCTCCAGACTGTCCTGTGTAAAAAATAGGAATAAAGGGGGTTTGTTA	712					
Qy	138	TTATTTTACTGATGTATATAATAATTTGTATAAGAAAATG	179					
Db	713	TTATTTTACTGATGTATATAATAATTTGTATAAGAAAATG	754					

Title: US-09-834-271A-26
Perfect score: 185
Sequence: 1 ggcctaaggcctcaatc.....tgtataaagaatggagctc 185
Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0
Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcto:*
9: gb_est1:*
10: gb_est2:*
11: gb_hct:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vil:*
28: gb_gss1:*
29: gb_gss2:*

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.
OM nucleic - nucleic search, using sw model
Run on: March 19, 2004, 01:21:15 ; Search time 1837 Seconds
(without alignments)
3007.351 Million cell updates/sec

5	39.2	21.2	699	28	BZ072552
6	38.6	20.9	1038	29	CN801TL7
7	38.4	20.8	700	28	BH922535
8	38.2	20.6	1201	13	BX356191
9	37.6	20.3	392	10	AW292419
10	37.6	20.3	461	9	AW001963
11	37.6	20.3	462	12	BQ000230
12	37.6	20.3	467	9	AI989386
13	37.6	20.3	475	9	AI697642
14	37.6	20.3	564	9	AI479617
15	37.6	20.3	700	9	AU253100
16	37.2	20.1	868	28	AZ535134
17	37.2	20.1	887	28	AZ546631
18	37.2	20.1	1201	13	BX385059
19	37.2	20.1	1201	13	BX419991
20	37	20.0	509	10	BE157233
21	37	20.0	686	28	BZ028517
22	37	20.0	727	29	CC920927
23	37	20.0	1010	10	BE614304
24	37	20.0	1304	28	CC317939
25	36.8	19.9	523	28	B25666
26	36.8	19.9	1201	9	AL543368
27	36.6	19.8	403	9	AI499797
28	36.6	19.8	876	28	AZ691916
29	36.6	19.8	879	28	AZ672806
30	36.6	19.8	879	28	AZ672806
31	36.6	19.8	1201	9	AL513973
32	36.4	19.7	505	28	AQ460935
33	36.4	19.7	1124	13	BX436382
34	36.4	19.7	1200	13	BX426107
35	36.2	19.6	867	28	BZ446197
36	36.2	19.6	916	14	CA471136
37	36.2	19.6	928	29	CNS000KY
38	36.2	19.6	1101	29	CNS000HP5
39	36	19.5	565	12	BM153720
40	36	19.5	986	29	CNS000EH
41	36	19.5	1040	29	CNS01820
42	36	19.5	1062	28	CC305992
43	35.8	19.4	418	12	BM959532
44	35.8	19.4	453	29	CE516519
45	35.8	19.4	894	9	AL543011

ALIGNMENTS

RESULT 1
CA819917
LOCUS CA819917
DEFINITION sau82a05.Y1 Gm-cl048 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl048-2314 5', mRNA sequence.
ACCESSION CA819917
VERSION CA819917.1
KEYWORDS GI:26268854
SOURCE EST
ORGANISM Glycine max (soybean)
Glycine max
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
1 (bases 1 to 427)
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,
Khanna,A., Bolla,B., Marita,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	24.3	427	14	CA819917 sau82a05.
2	40.6	21.9	1201	9	AL544700 AL544700
3	40.4	21.8	735	29	CE808066 tigr-gss-
4	40.2	21.7	1201	13	EX446359 EX446359


```

/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match          21.8%; Score 40.4; DB 29; Length 735;
Best Local Similarity 53.9%; Pred. No. 8.9;
Matches 83; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 30 AGAAGAAGAGAGACATAAAATACCTTGTCTGCATCAGACAGGATTTTATGCT 89
Db 15 ACAAAAACAAACAAACAAATTCCTGTTTTCCTCTTTTCTTATTTCTTTT 74
QY 90 CTCAGACTGTCGCTGTGTAATAAAGGATTAAGGGGGTTCACATTTTACTGA 149
Db 75 TTTTGTGTTTCTCTCTGTTGTAATAAATGAATGAATGAACACCATTTGTTT 134
QY 150 TATGATATATATAATTTGATTAAGAAATGGAGC 183
Db 135 GTTATTATTATTATTAGACACGAGAGAGGAGC 168

RESULT 4
BX446359/c
LOCUS
DEFINITION
BX446359 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA003ZG04
3-PRIME mRNA sequence.
ACCESSION
BX446359
VERSION
BX446359.1 GI:31023719
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2980.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBA003ZG04P1&cluster=2980.r. Contact :
Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBA003ZG04P1.

FEATURES
source
1..1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA003ZG04"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and ScaRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match          21.7%; Score 40.2; DB 13; Length 1201;
Best Local Similarity 40.5%; Pred. No. 8.4;
Matches 70; Conservative 27; Mismatches 76; Indels 0; Gaps 0;

QY 5 TTAAGCGCTCGCAATCGATTGTTGAGAAAGAGAGACCAATAAAATACCTTCTGCT 64
Db 1024 TGAARAWTTCGGTTCGGTGGTGGGGGAGAAACCCSYGTCGTGAACMCCMCACWNA 965

```

```

65 CATCAGACAGGGTATTTTATCTGTCCAGACTGTCGCTGTGTGTATATAAAGGAATAA 124
Db 964 CACGAAACGATGKTKTGTCCTGTGGCGWTKTTTTTCTGKGRADAAAARAAWW 905
QY 125 AGGGGGTTCACATTTTACTGATATGTATATAATTTTGTATAAGAAA 177
Db 904 CTGGGGCTGYTYYTGTGABAAAAAANAANAANAANAANAANAANA 852

RESULT 5
BX072552
LOCUS
DEFINITION
BX072552 b1 B.oleracea002 Brassica oleracea genomic, genomic survey
sequence.
ACCESSION
BX072552
VERSION
BX072552.1 GI:23690695
KEYWORDS
GSS.
SOURCE
Brassica oleracea
ORGANISM
Brassica oleracea
REFERENCE
Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,
Nash, W., Rabinowicz, P.D. and Wilson, R.K.
Whole genome shotgun reads from Brassica oleracea
Unpublished (2002)
JOURNAL
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: lkf46 row: b column: 09
Seq primer: -21UPPOT forward
Class: shotgun
High quality sequence start: 17
High quality sequence stop: 551.
Location/Qualifiers
1..699
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/note="Vector: POTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea T01000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

ORIGIN
Query Match          21.2%; Score 39.2; DB 28; Length 699;
Best Local Similarity 53.2%; Pred. No. 17;
Matches 83; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 22 ATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGTCTGTCATCAGACGGTATTT 81
Db 328 ATTGTTAAAAATTATATATAATAAAATAAATAATTAATTATTAATTTATGTTAT 387
QY 82 TTTATGCTGCCAGACTGTCGCTGTGTATAAANAAGGAATAAAGGGGGTTCACATTAT 141
Db 388 TTAATATTGTGAAGACTAGATTTTGTAGAAACTGAGGAAATACCATGTTATGTCAT 447
QY 142 TTTACTGATATGATATATAATTTGTATAAGAAA 177
Db 448 TTTATAAATGAATTAAGATTATAGATTATTAACA 483

RESULT 6
CNS01TL7
LOCUS
DEFINITION
CNS01TL7 1038 bp DNA linear GSS 01-SEP-2000
Tetradon nigroviridis genome survey sequence T7 end of clone

```


Email: cgabbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 696 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 448.
Location/Qualifiers

FEATURES

source

1. .461
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2513953"
/dev_stage="fetal"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares_thymus_NHFrh"
/note="Organ: thymus, pooled; Vector: p773D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo (dT) primer [5',
TGTTACCAATCTGAAGTGGGCGCGCAACGTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 20.3%; Score 37.6; DB 9; Length 461;
Best Local Similarity 72.1%; Pred. No. 48;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 111 AAAAAAGGAATAAGGGGGTTGACATATTTTACTGATATGATATATAATTTGTAT 170
Db 176 AAGAGAGGAGGAGAACTAAGTGTGTCATTAATGACAGATAAGTAAATAATTTGCTT 235
QY 171 AAGAAAT 178
Db 236 AAGAAAT 243

RESULT 11

BQ000230/c

LOCUS

DEFINITION UI-H-D10-avp-j-04-0-UI.s1 NCI_CGAP_D10 Homo sapiens cDNA clone
IMAGE:5882595 3', mRNA sequence.

ACCESSION

BQ000230

VERSION

BQ000230.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
The following repetitive elements were found in this cDNA
sequence: 433-454, >AT-rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

Location/Qualifiers
1. .462
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:5882595"
/tissue_type="Lung focal Fibrosis"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_D10"
/note="Organ: Lung; Vector: p773D-Pac (Pharmacia) with a
modified polylinker; Site 1: Eco RI; Site 2: Not I;
NCI_CGAP_D10 is a cDNA library containing the following
tissue(s): A pool of Lung focal Fibrosis. The library was
constructed according to Bonaldo, Lemmon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoRI
adaptor, digested with Not I, and cloned directionally
into p773D-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
ATACGGCGTC.
TAG TISSUE=Lung with fibrosis
TAG LIB=UI-H-D10
TAG_SEQ=ATACGGCGTC"

ORIGIN

Query Match 20.3%; Score 37.6; DB 12; Length 462;
Best Local Similarity 72.1%; Pred. NO. 48;
Matches 49; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 111 AAAAAAGGAATAAGGGGGTTGACATATTTTACTGATATGATATAATAATTGTAT 170
Db 270 AAGAGAGGAGGAGAACTAAGTGTGTCATTAATGACAGATAAGTAAATAATTTGCTT 211
QY 171 AAGAAAT 178
Db 210 AAGAAAT 203

RESULT 12

AI989386

LOCUS

DEFINITION

wt8e12.x1 Soares_thymus_NHFrh Homo sapiens cDNA clone

IMAGE:2513806 3', mRNA sequence.

ACCESSION

AI989386

VERSION

AI989386.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40UP from Gibco

High quality sequence stop: 442.

Location/Qualifiers

1. .467

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2513806"

/dev_stage="fetal"

/lab_host="Soares_thymus_NHFrh"

/clone_lib="Soares_thymus_NHFrh"

/note="Organ: thymus, pooled; Vector: p773D-Pac

(Pharmacia) with a modified polylinker; Site 1: Not I;

Site 2: Eco RI; 1st strand cDNA was primed with a Not I -

AI989386 467 bp mRNA linear EST 08-SEP-1999
wt8e12.x1 Soares_thymus_NHFrh Homo sapiens cDNA clone
IMAGE:2513806 3', mRNA sequence.

ACCESSION

AI989386

VERSION

AI989386.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40UP from Gibco

High quality sequence stop: 442.

Location/Qualifiers

1. .467

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2513806"

/dev_stage="fetal"

/lab_host="Soares_thymus_NHFrh"

/clone_lib="Soares_thymus_NHFrh"

/note="Organ: thymus, pooled; Vector: p773D-Pac

(Pharmacia) with a modified polylinker; Site 1: Not I;

Site 2: Eco RI; 1st strand cDNA was primed with a Not I -

oligo(dT) primer [5'
TGTACATCTGACGTGGCGCGCCGACAGTGTGTGTGTGTGTGT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo.

Search completed: March 19, 2004, 03:16:12
Job time : 1842 secs

Search completed: March 19, 2004, 03:16:12
Job time : 1842 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 02:14:41 ; Search time 745.5 Seconds
(without alignments)
917.966 Million cell updates/sec

Title: US-09-834-271A-26
Perfect score: 185
Sequence: 1 ggccttaaggccctgaatc.....tgtataagaaatggagctc 185

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	185	100.0	185	10	US-09-834-271A-26
2	185	100.0	185	14	US-10-406-025-5
3	183.4	99.1	185	10	US-09-834-271A-27
4	183.4	99.1	185	14	US-10-406-025-6
5	178.6	96.5	185	10	US-09-834-271A-25
6	178.6	96.5	185	14	US-10-406-025-7
7	159.6	86.3	2166	12	US-10-344-231-17
8	159.6	86.3	2166	12	US-10-363-332A-17
9	159.6	86.3	2267	12	US-10-344-231-18
10	159.6	86.3	2267	12	US-10-363-332A-18
11	159.6	86.3	2588	12	US-10-344-231-20
12	159.6	86.3	2588	12	US-10-363-332A-20
13	155.6	84.1	2084	9	US-09-769-864-58
14	155.6	84.1	2084	9	US-09-854-346-9
15	155.6	84.1	2084	9	US-09-918-543-9

16	155.6	84.1	2084	10	US-09-925-576C-9	Sequence 9, Appli
17	155.6	84.1	2084	12	US-10-665-667-58	Sequence 58, Appli
18	155.6	84.1	2084	12	US-10-644-187-3	Sequence 3, Appli
19	155.6	84.1	2084	14	US-10-184-771-3	Sequence 3, Appli
20	155.6	84.1	2084	14	US-10-186-042-3	Sequence 3, Appli
21	155.6	84.1	2604	14	US-10-146-327-5	Sequence 5, Appli
22	155.6	84.1	2604	14	US-09-736-116-60	Sequence 60, Appli
23	88.6	47.5	6837	10	US-09-928-847B-49	Sequence 49, Appli
24	44	23.8	44	10	US-09-834-271A-19	Sequence 19, Appli
25	39	21.1	42	10	US-09-834-271A-9	Sequence 10, Appli
26	38.4	20.8	43	10	US-09-834-271A-10	Sequence 10, Appli
C 27	37.6	20.3	2141	15	US-10-104-047-1737	Sequence 1737, Ap
C 28	37	20.0	3673778	14	US-10-312-841-2	Sequence 2, Appli
C 29	36	19.5	28588	9	US-09-764-887-399	Sequence 399, Appli
C 30	36	19.5	28588	14	US-10-073-961-339	Sequence 399, Appli
C 31	35.6	19.2	179	9	US-09-783-590-7881	Sequence 7881, Ap
C 32	35.2	19.0	317	15	US-10-125-968-1260	Sequence 1260, Ap
C 33	34.8	18.8	3673778	14	US-10-312-841-1	Sequence 1, Appli
C 34	34.6	18.7	14147	12	US-10-221-744A-470	Sequence 470, App
C 35	34.6	18.7	14147	14	US-10-172-086-52	Sequence 52, Appli
C 36	34.4	18.6	383	9	US-09-969-347-299	Sequence 299, Appli
C 37	34.2	18.5	9121	9	US-09-070-927A-221	Sequence 221, Appli
C 38	34.2	18.5	3673778	14	US-10-312-841-1	Sequence 1, Appli
C 39	33.8	18.3	347	12	US-10-424-599-69129	Sequence 69129, A
C 40	33.8	18.3	474	10	US-09-764-891-1477	Sequence 1477, Ap
C 41	33.8	18.3	158001	16	US-10-211-179-11	GENERAL INFORMATI
C 42	33.6	18.2	512	15	US-10-027-632-15736	Sequence 15736, A
C 43	33.6	18.2	827	15	US-10-027-632-145937	Sequence 145937,
C 44	33.6	18.2	19380	12	US-10-221-613-399	Sequence 389, App
C 45	33.6	18.2	185695	14	US-10-020-141-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1
US-09-834-271A-26
; Sequence 26, Application US/09834271A
; Publication No. US20030170876A1
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/834,271A
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US/09/258,377
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/031,442
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
; US-09-834-271A-26

Query Match	100.0%	Score 185;	DB 10;	Length 185;
Best Local Similarity	100.0%;	Pred. No. 1.3e-41;		
Matches 185;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GGCTTAAAGGCGCTGCAATCGATTGTTGAGAAAGAGAGACCAATAAAATACCTTGT	60	
Db	1	GGCTTAAAGGCGCTGCAATCGATTGTTGAGAAAGAGAGACCAATAAAATACCTTGT	60	
Qy	61	CTGTCTATCAGACAGGGTATTTTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAAAGGA	120	
Db	61	CTGTCTATCAGACAGGGTATTTTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAAAGGA	120	
Qy	121	ATAAAGGGGGTTCACATTAATTTACTGATATGATATATAATTTGTATAAGAAATGG	180	

Db 121 ATAAAGGGGGTGCATTTACTGATATGATATAATTTGTTATAGAAATGG 180
181 AGCTC 185
181 AGCTC 185

RESULT 2
US-10-406-025-5
; Sequence 5, Application US/10406025
; Publication No. US20030186380A1
; GENERAL INFORMATION:
; APPLICANT: No. US20030186380A1; Thomas, Michael D.
; APPLICANT: Thomas, Michael D.
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Methods for producing secreted polypeptides having L-asparaginase
; FILE REFERENCE: 10289.200-US
; CURRENT APPLICATION NUMBER: US/10/406,025
; PRIOR FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/369,192
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-10-406-025-5

Query Match 100.0%; Score 185; DB 14; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.3e-41;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGGCTGCAATCGATTGTTGAGAAAAGAGAGACCATATAAATACCTTGT 60
DB 1 GGCCTTAAGGGGCTGCAATCGATTGTTGAGAAAAGAGAGACCATATAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCGTGTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCGTGTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTGCATTTACTGATATGATATAATTTGTTATAGAAATGG 180
DB 121 ATAAAGGGGGTGCATTTACTGATATGATATAATTTGTTATAGAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 3
US-09-834-271A-27
; Sequence 27, Application US/09834271A
; Publication No. US20030170876A1
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/834,271A
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US/09/258,377
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/031,442
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 185

; TYPE: DNA
; ORGANISM: Bacillus
US-09-834-271A-27

Query Match 99.1%; Score 183.4; DB 10; Length 185;
Best Local Similarity 99.5%; Pred. No. 3.5e-41;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGGCTGCAATCGATTGTTGAGAAAAGAGAGACCATATAAATACCTTGT 60
DB 1 GGCCTTAAGGGGCTGCAATCGATTGTTGAGAAAAGAGAGACCATATAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCGTGTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCGTGTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTGCATTTACTGATATGATATAATTTGTTATAGAAATGG 180
DB 121 ATAAAGGGGGTGCATTTACTGATATGATATAATTTGTTATAGAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 4
US-10-406-025-6
; Sequence 6, Application US/10406025
; Publication No. US20030186380A1
; GENERAL INFORMATION:
; APPLICANT: No. US20030186380A1; Thomas, Michael D.
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods for producing secreted polypeptides having L-asparagina
; FILE REFERENCE: 10289.200-US
; CURRENT APPLICATION NUMBER: US/10/406,025
; PRIOR FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/369,192
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-10-406-025-6

Query Match 99.1%; Score 183.4; DB 14; Length 185;
Best Local Similarity 99.5%; Pred. No. 3.5e-41;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGGCTGCAATCGATTGTTGAGAAAAGAGAGACCATATAAATACCTTGT 60
DB 1 GGCCTTAAGGGGCTGCAATCGATTGTTGAGAAAAGAGAGACCATATAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCGTGTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCGTGTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTGCATTTACTGATATGATATAATTTGTTATAGAAATGG 180
DB 121 ATAAAGGGGGTGCATTTACTGATATGATATAATTTGTTATAGAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 5
US-09-834-271A-25
; Sequence 25, Application US/09834271A
; Publication No. US20030170876A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Widner, William
/ APPLICANT: Sloma, Alan
/ APPLICANT: Thomas, Michael D.
/ TITLE OF INVENTION: Methods For Producing A polypeptide In a
/ TITLE OF INVENTION: Bacillus Cell
/ FILE REFERENCE: 5455.200-US
/ CURRENT APPLICATION NUMBER: US/09/834,271A
/ CURRENT FILING DATE: 2001-04-12
/ PRIOR APPLICATION NUMBER: US/09/258,377
/ PRIOR FILING DATE: 1999-02-26
/ PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/031,442
/ PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-26
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 25
/ LENGTH: 185
/ TYPE: DNA
/ ORGANISM: Bacillus
US-09-834-271A-25

Query Match          96.5%; Score 178.6; DB 10; Length 185;
Best Local Similarity 97.8%; Pred. No. 7.6e-40;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTCGAATCGATTGTTGAGAAAAGAGAGACCATAAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTCGAATCGATTGTTGAGAAAAGAGAGACCATAAAAATACCTTGT 60

QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAAGGA 120

QY 121 ATAAAGGGGGTGTGACATTTTACTGATATGTAATAATAATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTGTGATTTTACTGATATGTAATAATAATTTGTATAGAAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 6
US-10-406-025-7
/ Sequence 7, Application US/10406025
/ Publication No. US20030186380A1
/ GENERAL INFORMATION:
/ APPLICANT: No. US20030186380A1ozymes Biotech, Inc.
/ APPLICANT: Sloma, Alan
/ TITLE OF INVENTION: Methods for producing secreted polypeptides having L-asparaginase
/ TITLE OF INVENTION: activity
/ FILE REFERENCE: 10289.200-US
/ CURRENT APPLICATION NUMBER: US/10/406,025
/ CURRENT FILING DATE: 2003-04-01
/ PRIOR APPLICATION NUMBER: US 60/369,192
/ PRIOR FILING DATE: 2002-04-01
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 7
/ LENGTH: 185
/ TYPE: DNA
/ ORGANISM: Bacillus
US-10-406-025-7

Query Match          96.5%; Score 178.6; DB 14; Length 185;
Best Local Similarity 97.8%; Pred. No. 7.6e-40;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTCGAATCGATTGTTGAGAAAAGAGAGACCATAAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTCGAATCGATTGTTGAGAAAAGAGAGACCATAAAAATACCTTGT 60
```

```
QY 61 CTCTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAAGGA 120
DB 61 CTCTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAAGGA 120

QY 121 ATAAAGGGGGTGTGACATTTTACTGATATGTAATAATAATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTGTGATTTTACTGATATGTAATAATAATTTGTATAGAAAATGG 180

QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 7
US-10-344-231-17
/ Sequence 17, Application US/10344231
/ Publication No. US20040038845A1
/ GENERAL INFORMATION:
/ APPLICANT: Pedersen, Poul
/ APPLICANT: Roggen, Erwin Ludo
/ TITLE OF INVENTION: Improved production of proteases with inhibitors
/ FILE REFERENCE: 10064.204-US
/ CURRENT APPLICATION NUMBER: US/10/344,231
/ CURRENT FILING DATE: 2003-02-10
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 17
/ LENGTH: 2166
/ TYPE: DNA
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic
US-10-344-231-17

Query Match          86.3%; Score 159.6; DB 12; Length 2166;
Best Local Similarity 97.6%; Pred. No. 3.5e-34;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGTCTGTCATCAGACA 73
DB 172 TGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGTCTGTCATCAGACA 231

QY 74 GGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAAGGAATAAGGGGGTT 133
DB 232 GGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAAGGAATAAGGGGGTT 291

QY 134 GACATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 179
DB 292 GTTATTTTACTGATATGTAATAATAATTTGTATAGAAAATG 337

RESULT 8
US-10-363-332A-17
/ Sequence 17, Application US/10363332A
/ Publication No. US20040038375A1
/ GENERAL INFORMATION:
/ APPLICANT: Pedersen, Poul
/ APPLICANT: No. US20040038375A1regard-Madsen, Mads
/ TITLE OF INVENTION: Method For Screening Highly Active Proteases and Inhibitors
/ FILE REFERENCE: 10080.204-US
/ CURRENT APPLICATION NUMBER: US/10/363,332A
/ CURRENT FILING DATE: 2003-05-27
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 17
/ LENGTH: 2166
/ TYPE: DNA
/ ORGANISM: Artificial
/ FEATURE:
/ OTHER INFORMATION: Construct A
US-10-363-332A-17

Query Match          86.3%; Score 159.6; DB 12; Length 2166;
```

Best Local Similarity 97.6%; Pred. No. 3.5e-34;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 TCGAATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 73
DB 172 TCGAATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTAAATAAGGAATAAAGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTAAATAAGGAATAAAGGGGGTT 291

QY 134 GACATTTATTTACTGATATGTAATAATTAATTTGTAAGAAATG 179
DB 292 GTTATTTATTTACTGATATGTAATAATTAATTTGTAAGAAATG 337

RESULT 9

US-10-344-231-18
; Sequence 18, Application US/10344231
; Publication No. US20040038845A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; TITLE OF INVENTION: Improved production of proteases with inhibitors
; FILE REFERENCE: 10064.204-US
; CURRENT APPLICATION NUMBER: US/10/344,231
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 2267
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-344-231-18

Query Match 86.3%; Score 159.6; DB 12; Length 2267;
Best Local Similarity 97.6%; Pred. No. 3.6e-34;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 TCGAATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 73
DB 172 TCGAATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTAAATAAGGAATAAAGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTAAATAAGGAATAAAGGGGGTT 291

QY 134 GACATTTATTTACTGATATGTAATAATTAATTTGTAAGAAATG 179
DB 292 GTTATTTATTTACTGATATGTAATAATTAATTTGTAAGAAATG 337

RESULT 10

US-10-363-332A-18
; Sequence 18, Application US/10363332A
; Publication No. US20040038375A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; TITLE OF INVENTION: Method For Screening Highly Active Proteases and Inhibitors
; FILE REFERENCE: 10080.204-US
; CURRENT APPLICATION NUMBER: US/10/363,332A
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 2267
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Construct B

US-10-363-332A-18

Query Match 86.3%; Score 159.6; DB 12; Length 2267;
Best Local Similarity 97.6%; Pred. No. 3.6e-34;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 TCGAATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 73
DB 172 TCGAATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTAAATAAGGAATAAAGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTAAATAAGGAATAAAGGGGGTT 291

QY 134 GACATTTATTTACTGATATGTAATAATTAATTTGTAAGAAATG 179
DB 292 GTTATTTATTTACTGATATGTAATAATTAATTTGTAAGAAATG 337

RESULT 11

US-10-344-231-20
; Sequence 20, Application US/10344231
; Publication No. US20040038845A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; TITLE OF INVENTION: Improved production of proteases with inhibitors
; FILE REFERENCE: 10064.204-US
; CURRENT APPLICATION NUMBER: US/10/344,231
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 2588
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc feature
; LOCATION: (2070)..(2070)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2077)..(2077)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2084)..(2084)
; OTHER INFORMATION: n is a, c, g, or t
US-10-344-231-20

Query Match 86.3%; Score 159.6; DB 12; Length 2588;
Best Local Similarity 97.6%; Pred. No. 3.8e-34;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 TCGAATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 73
DB 172 TCGAATCGATTGTTTGAGAAAGAGAGACCATATAAAATACCTTGTCTGTCTGCATCAGACA 231

QY 74 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTAAATAAGGAATAAAGGGGGTT 133
DB 232 GGGTATTTTATGCTGCCAGACTGTCGGCTGTGTAAATAAGGAATAAAGGGGGTT 291

QY 134 GACATTTATTTACTGATATGTAATAATTAATTTGTAAGAAATG 179
DB 292 GTTATTTATTTACTGATATGTAATAATTAATTTGTAAGAAATG 337

RESULT 12

US-10-363-332A-20
; Sequence 20, Application US/10363332A
; Publication No. US20040038375A1

```

; GENERAL INFORMATION:
; APPLICANT: Pedersen, Poul
; APPLICANT: No. US2004003875Alreggaard-Madsen, Mads
; TITLE OF INVENTION: Method For Screening Highly Active Proteases and Inhibitors
; FILE REFERENCE: 10080.204-US
; CURRENT APPLICATION NUMBER: US/10/363,332A
; CURRENT FILING DATE: 2003-05-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 20
; LENGTH: 2598
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Construct D
; NAME/KEY: misc feature
; LOCATION: (2070)..(2070)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2077)..(2077)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2084)..(2084)
; OTHER INFORMATION: n is a, c, g, or t
; OTHER INFORMATION: n is a, c, g, or t
; US-10-363-332A-20

Query Match      86.3%; Score 159.6; DB 12; Length 2598;
Best Local Similarity 97.6%; Pred. No. 3.8e-34; Indels 0; Gaps 0;
Matches 162; Conservative 0; Mismatches 4;

QY 14 TCGAATCGATTGTTGAGAAAAGAGAACCAACCAATAAAATACCTTCTGTCTCATCAGACA 73
DB 172 TCGAATCGATTGTTGAGAAAAGAGAACCAACCAATAAAATACCTTCTGTCTCATCAGACA 231

QY 74 GGGTATTTTATGCTGCTCCAGACTGCCGCTGTGTAATAAAGGAATAAAGGGGGTT 133
DB 232 GGGTATTTTATGCTGCTCCAGACTGCCGCTGTGTAATAAAGGAATAAAGGGGGTT 291

QY 134 GACATTATTTTACTGATGATGATATATATATATATATATATATATATATATATATG 179
DB 292 GTTATTTATTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 337

RESULT 13
US-09-769-864-58
; Sequence 58, Application US/09769864
; Patent No. US20010039253A1
; GENERAL INFORMATION:
; APPLICANT: Borchert, Torben V.
; APPLICANT: Svendsen, Allan
; APPLICANT: Andersen, Carsten
; APPLICANT: Nielsen, Bjarne
; APPLICANT: Nissen, Torben L.
; APPLICANT: Kiaerulff, Soren
; TITLE OF INVENTION: Alpha-Amulase Mutants
; FILE REFERENCE: 5368.200-US
; CURRENT APPLICATION NUMBER: US/09/769,864
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/183,412
; PRIOR FILING DATE: 1998-10-30
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (343)....(1794)
; OTHER INFORMATION: BAN
; US-09-769-864-58

Query Match      84.1%; Score 155.6; DB 9; Length 2084;
Best Local Similarity 97.5%; Pred. No. 4.5e-33;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTCTGTCTCATCAGACGGGT 77
DB 73 ATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTCTGTCTCATCAGACGGGT 132

QY 78 ATTTTATGCTGCTCCAGACTGCCGCTGTGTAATAAAGGAATAAAGGGGGTTGACA 137
DB 133 ATTTTATGCTGCTCCAGACTGCCGCTGTGTAATAAAGGAATAAAGGGGGTTGTTA 192

QY 138 TTATTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 179
DB 193 TTATTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 234

RESULT 14
US-09-854-346-9
; Sequence 9, Application US/09854346
; Patent No. US20020068352A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020068352A1ozymes A/S
; APPLICANT: Svendsen, Allan
; APPLICANT: Jorgensen, Christel Thea
; APPLICANT: Nielsen, Bjarne Romfeldt
; TITLE OF INVENTION: Alpha-amyase variants with altered 1.6 activity
; FILE REFERENCE: 6140.200-US
; CURRENT APPLICATION NUMBER: US/09/854,346
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 9
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (343)....(1794)
; OTHER INFORMATION: BAN
; US-09-854-346-9

Query Match      84.1%; Score 155.6; DB 9; Length 2084;
Best Local Similarity 97.5%; Pred. No. 4.5e-33;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTCTGTCTCATCAGACGGGT 77
DB 73 ATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTCTGTCTCATCAGACGGGT 132

QY 78 ATTTTATGCTGCTCCAGACTGCCGCTGTGTAATAAAGGAATAAAGGGGGTTGACA 137
DB 133 ATTTTATGCTGCTCCAGACTGCCGCTGTGTAATAAAGGAATAAAGGGGGTTGTTA 192

QY 138 TTATTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 179
DB 193 TTATTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 234

RESULT 15
US-09-918-543-9
; Sequence 9, Application US/09918543
; Patent No. US20020155574A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020155574A1ozymes A/S
; APPLICANT: Thisted, Thomas
; APPLICANT: Kiaerulff, Soren
; APPLICANT: Andersen, Carsten
; APPLICANT: Fuglsang, Claus Crone
; TITLE OF INVENTION: Alpha-amyase mutants with altered properties
; FILE REFERENCE: 10062.200-US
; CURRENT APPLICATION NUMBER: US/09/918,543
```

Mon Mar 22 09:21:45 2004

```
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (343)..(1794)
; OTHER INFORMATION:
US-09-918-543-9

Query Match      84.1%; Score 155.6; DB 9; Length 2084;
Best Local Similarity 97.5%; Pred. No. 4.5e-33;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      18 ATCGATTGTTGAGAAAAGAGAGACCATAAAATACCTTGTCTGTCTCATCAGACAGGGT 77
Db      73 ATCGATTGTTGAGAAAAGAGAGACCATAAAATACCTTGTCTGTCTCATCAGACAGGGT 132

QY      78 ATTTTATGCTGTCCGACTGTCCGCTGTGTAAAAAAGGAATAAAGGGGGGTGACA 137
Db      133 ATTTTATGCTGTCCGACTGTCCGCTGTGTAAAAAATAGGATTAAGGGGGGTGCTTA 192

QY      138 TTATTTTACTGATATGTAATAATAATTGTAAGAAATG 179
Db      193 TTATTTTACTGATATGTAATAATAATTGTAAGAAATG 234

Search completed: March 19, 2004, 05:16:19
Job time : 753.5 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 01:23:55 ; Search time 44 Seconds
(without alignments)
2333.316 Million cell updates/sec

Title: US-09-834-271A-26
Perfect score: 185
Sequence: 1 gsccttaaggccctgcaatc.....tgtatagaataaggagctc 185

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A-COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B-COMB.seq: *
3: /cgn2_6/prodata/2/ina/5A-COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B-COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCITUS-COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185	100.0	185	3	US-09-384-305-21
2	185	100.0	185	3	US-09-258-377-26
3	183.4	99.1	185	3	US-09-384-305-22
4	183.4	99.1	185	3	US-09-258-377-27
5	178.6	96.5	185	3	US-09-384-305-20
6	178.6	96.5	185	3	US-09-258-377-25
7	159.6	86.3	10216	2	US-08-875-154-1
8	155.6	84.1	162	1	US-08-434-255-24
9	155.6	84.1	162	1	US-08-459-967-24
10	155.6	84.1	162	1	US-08-460-327-24
11	155.6	84.1	162	1	US-08-459-871-24
12	155.6	84.1	2084	1	US-08-720-899-3
13	155.6	84.1	2084	1	US-08-459-610-3
14	155.6	84.1	2084	2	US-08-343-804-3
15	155.6	84.1	2084	2	US-08-687-399-3
16	155.6	84.1	2084	2	US-08-600-908A-3
17	155.6	84.1	2084	3	US-08-683-838A-3
18	155.6	84.1	2084	3	US-09-182-859-3
19	155.6	84.1	2084	3	US-09-170-670-13
20	155.6	84.1	2084	3	US-09-193-068-31
21	155.6	84.1	2084	3	US-09-183-412-58
22	155.6	84.1	2084	4	US-09-290-734-13
23	155.6	84.1	2084	4	US-09-672-459-3
24	155.6	84.1	2084	4	US-09-636-252A-3
25	155.6	84.1	2084	4	US-09-545-586-13
26	155.6	84.1	2084	4	US-10-186-042-3
27	155.6	84.1	2084	4	US-09-769-864-58

RESULT 1
US-09-384-305-21
; Sequence 21, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:
; APPLICANT: Michael D. Thomas
; APPLICANT: Kimberly M. Brown
; TITLE OF INVENTION: Polypeptides Having Pectin
; TITLE OF INVENTION: Acetyltransferase Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5952.000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-384-305-21

ALIGNMENTS

Query Match 100.0%; Score 185; DB 3; Length 185;
Best Local Similarity 100.0%; Pred. No. 9.7e-43;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGCCTTAAGGCGCTGCATCGATTGTTTGAGAAAGAGACCATATAAAATACCTTGT	60
Db	1	GGCCTTAAGGCGCTGCATCGATTGTTTGAGAAAGAGACCATATAAAATACCTTGT	60
QY	61	CTGTCAATCAGACAGGGTATTTTATGCTGTCAGACTGCCCTGTGTAATAAAGGA	120
Db	61	CTGTCAATCAGACAGGGTATTTTATGCTGTCAGACTGCCCTGTGTAATAAAGGA	120
QY	121	ATAAAGGGGGTGGACATTTTACTGATATGATATATAATATGTAAGAAATGG	180
Db	121	ATAAAGGGGGTGGACATTTTACTGATATGATATATAATATGTAAGAAATGG	180
QY	181	AGCTC 185	
Db	181	AGCTC 185	

RESULT 2
US-09-258-377-26
; Sequence 26, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a

Sequence 3, Appli
Sequence 5, Appli
Sequence 15, Appli
Sequence 19, Appli
Sequence 37, Appli
Sequence 9, Appli
Sequence 25, Appli
Sequence 9, Appli
Sequence 28, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 26, Appli
Sequence 26, Appli
Sequence 26, Appli
Sequence 1, Appli
Sequence 8, Appli

```

; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/031,442
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-09-258-377-26

Query Match          100.0%; Score 185; DB 3; Length 185;
Best Local Similarity 100.0%; Pred. No. 9.7e-43;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTTGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 3
US-09-384-305-22
; Sequence 22, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:
; APPLICANT: Michael D. Thomas
; APPLICANT: Kimberly M. Brown
; TITLE OF INVENTION: Polypeptides Having Pectin
; TITLE OF INVENTION: Acetyltransferase Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5952.000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-384-305-22

Query Match          99.1%; Score 183.4; DB 3; Length 185;
Best Local Similarity 99.5%; Pred. No. 2.7e-42;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTTGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 4
US-09-258-377-27
; Sequence 27, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Thomas, Michael D.
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/031,442
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-09-258-377-27

Query Match          99.1%; Score 183.4; DB 3; Length 185;
Best Local Similarity 99.5%; Pred. No. 2.7e-42;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTTGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 5
US-09-384-305-20
; Sequence 20, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:
; APPLICANT: Michael D. Thomas
; APPLICANT: Kimberly M. Brown
; TITLE OF INVENTION: Polypeptides Having Pectin
; TITLE OF INVENTION: Acetyltransferase Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5952.000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-384-305-20

Query Match          96.5%; Score 178.6; DB 3; Length 185;
Best Local Similarity 97.8%; Pred. No. 5.8e-41;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
```

```

DB 181 AGCTC 185

RESULT 4
US-09-258-377-27
; Sequence 27, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Thomas, Michael D.
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/031,442
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-09-258-377-27

Query Match          99.1%; Score 183.4; DB 3; Length 185;
Best Local Similarity 99.5%; Pred. No. 2.7e-42;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTTGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTGACATTATTTTACTGATATGATATATAATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 5
US-09-384-305-20
; Sequence 20, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:
; APPLICANT: Michael D. Thomas
; APPLICANT: Kimberly M. Brown
; TITLE OF INVENTION: Polypeptides Having Pectin
; TITLE OF INVENTION: Acetyltransferase Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5952.000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-384-305-20

Query Match          96.5%; Score 178.6; DB 3; Length 185;
Best Local Similarity 97.8%; Pred. No. 5.8e-41;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCCTGCAATCGATTGTTTGAGAAAAGAGAACCATATAAAATACCTTGT 60
```


Db 1 GGCTTAAGGCGCTCAATCGATTGTTTGAGAAAGAGAACCATATAAAATACCTTGT 60
QY 61 CTGTATCATCAGAGGTATTTTTATGCTGTCAGACTGTCGCTGTGTGTAAGAAAGGA 120
Db 61 CTGTATCATCAGAGGTATTTTTATGCTGTCAGACTGTCGCTGTGTGTAAGAAATAGGA 120
QY 121 ATAAAGGGGGTTGACATTATTTTACTGATATGATATATATATTTGTTATAGAAATGG 180
Db 121 ATAAAGGGGGTTGATTTATTTTACTGATATGTAATAATATTTGTTATAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 6

US-09-258-377-25
; Sequence 25, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455.200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/031,442
; EARLIER FILING DATE: 1998-02-26
; SOFTWARE: FastSeq for Windows Version 3.0
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 25
; LENGTH: 185
; TYPE: DNA
; ORGANISM: Bacillus
US-09-258-377-25

Query Match 96.5%; Score 178.6; DB 3; Length 185;

Best Local Similarity 97.8%; Pred. No. 5.9e-41;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCTTAAGGCGCTCAATCGATTGTTTGAGAAAGAGAACCATATAAAATACCTTGT 60
Db 1 GGCTTAAGGCGCTCAATCGATTGTTTGAGAAAGAGAACCATATAAAATACCTTGT 60
QY 61 CTGTATCATCAGAGGTATTTTTATGCTGTCAGACTGTCGCTGTGTGTAAGAAAGGA 120
Db 61 CTGTATCATCAGAGGTATTTTTATGCTGTCAGACTGTCGCTGTGTGTAAGAAATAGGA 120
QY 121 ATAAAGGGGGTTGACATTATTTTACTGATATGATATATATATTTGTTATAGAAATGG 180
Db 121 ATAAAGGGGGTTGATTTATTTTACTGATATGTAATAATATTTGTTATAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 7

US-08-875-154-1
; Sequence 1, Application US/08875154
; Patent No. 5882888
; GENERAL INFORMATION:
; APPLICANT: Jorgensen, Steven Truels
; TITLE OF INVENTION: DNA Integration By Transporation
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5882888 No. 5882888disk of No. 5882888th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA

ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/875,154
FILING DATE: 17-JUL-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4381.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10216 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "PMOL553"
US-08-875-154-1

Query Match 86.3%; Score 159.6; DB 2; Length 10216;
Best Local Similarity 97.6%; Pred. No. 2.5e-35;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 TGCATCGATTGTTTGAGAAAGAGAACCATATAAAATACCTTGTCTGTCATCAGACA 73
Db 2237 TGCATCGATTGTTTGAGAAAGAGAACCATATAAAATACCTTGTCTGTCATCAGACA 2296
QY 74 GGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAGAAAGGAATATAAGGGGGGTT 133
Db 2297 GGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAGAAATATAAGGGGGGTT 2356
QY 134 GACATTTTACTGATATGATATATAATTTGTTATAGAAATG 179
Db 2357 GTTATTTTACTGATATGTAATAATTTGTTATAGAAATG 2402

RESULT 8

US-08-434-255-24
; Sequence 24, Application US/08434255
; Patent No. 5621089
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dammann, Claus
; APPLICANT: Aaslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5621089c No. 5621089disk of No. 5621089th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,255
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Agiris Dr., Cheryl H.

REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 3764.400-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-434-255-24

Query Match 84.1%; Score 155.6; DB 1; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.4e-34;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGTCTGTCTATCAGACAGGGT 77
DB 1 ATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGTCTGTCTATCAGACAGGGT 60

QY 78 ATTTTACTGATGATGATAATAATATTTGTAAGAAAATG 179
DB 121 TTAATTTACTGATGATAATAATATTTGTAAGAAAATG 162

RESULT 9
US-08-459-967-24
Sequence 24, Application US/08459967
Patent No. 5622841
GENERAL INFORMATION:
APPLICANT: Sloma, Alan P.
APPLICANT: Outtrup, Helle
APPLICANT: Dammann, Claus
APPLICANT: Aaslytg, Dorrit
TITLE OF INVENTION: ALKALINE PROTEASE
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5622841o No. 5622841disk of No. 5622841th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,967
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/434,255
FILING DATE: 03-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 3764.400-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-459-967-24

Query Match 84.1%; Score 155.6; DB 1; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.4e-34;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGTCTGTCTATCAGACAGGGT 77
DB 1 ATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGTCTGTCTATCAGACAGGGT 60

QY 78 ATTTTACTGATGATGATAATAATATTTGTAAGAAAATG 179
DB 121 TTAATTTACTGATGATAATAATATTTGTAAGAAAATG 162

RESULT 10
US-08-460-327-24
Sequence 24, Application US/08460327
Patent No. 5622850
GENERAL INFORMATION:
APPLICANT: Sloma, Alan P.
APPLICANT: Outtrup, Helle
APPLICANT: Dammann, Claus
APPLICANT: Aaslytg, Dorrit
TITLE OF INVENTION: ALKALINE PROTEASE
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5622850o No. 5622850disk of No. 5622850th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,327
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/434,255
FILING DATE: 03-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Agis Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 3764.400-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-460-327-24

Query Match 84.1%; Score 155.6; DB 1; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.4e-34;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGTCTGTCTATCAGACAGGGT 77
DB 1 ATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGTCTGTCTATCAGACAGGGT 60

QY 78 ATTTTACTGATGATGATAATAATATTTGTAAGAAAATG 179
DB 121 TTAATTTACTGATGATAATAATATTTGTAAGAAAATG 162

Db 61 ATTTTATGCTGTCAGACTGCCGCTGTGTAAATAAGGAATAAGGGGGTTGTTA 120
Qy 138 TTATTTACTGATGTATATAATAATTTGTATAAGAAATG 179
Db 121 TTATTTACTGATGTATAATAATAATTTGTATAAGAAATG 162

RESULT 11

US-08-459-871-24
; Sequence 24, Application US/08459871
; Patent No. 5650326
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan P.
; APPLICANT: Outtrup, Helle
; APPLICANT: Dambmann, Claus
; APPLICANT: Aaslyng, Dorrit
; TITLE OF INVENTION: ALKALINE PROTEASE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5650326 No. 5650326disk of No. 5650326th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,871
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/434,255
; FILING DATE: 03-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Agis Dr., Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 3764.400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-459-871-24

Query Match 84.1%; Score 155.6; DB 1; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.4e-34;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 18 ATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGTCTGTCATCAGACAGGT 77
Db 1 ATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGTCTGTCATCAGACAGGT 60
Qy 78 ATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGAATAAAGGGGGTTGACA 137
Db 61 ATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGAATAAAGGGGGTTGTTA 120
Qy 138 TTATTTACTGATGTATATAATAATTTGTATAAGAAATG 179
Db 121 TTATTTACTGATGTATAATAATAATTTGTATAAGAAATG 162

RESULT 12

US-08-720-899-3
; Sequence 3, Application US/08720899

; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borcher, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5753460 No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowrey Dr. Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2084 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 250..1794
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 250..342
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 343..1791
US-08-720-899-3

Query Match 84.1%; Score 155.6; DB 1; Length 2084;
Best Local Similarity 97.5%; Pred. No. 2.3e-34;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 18 ATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGTCTGTCATCAGACAGGT 77
Db 73 ATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGTCTGTCATCAGACAGGT 132
Qy 78 ATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGAATAAAGGGGGTTGACA 137
Db 133 ATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGAATAAAGGGGGTTGTTA 192
Qy 138 TTATTTACTGATGTATATAATAATTTGTATAAGAAATG 179
Db 193 TTATTTACTGATGTATAATAATAATTTGTATAAGAAATG 234

RESULT 13

US-08-459-610-3
; Sequence 3, Application US/08459610
; Patent No. 5801043

TITLE OF INVENTION: A Combined Desizing and Bleaching
TITLE OF INVENTION: Process
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59283810 No. 59283810disk of No. 59283810th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM: disk
MEDIUM TYPE: Floppy
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,399
FILING DATE:
CLASSIFICATION: 008
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4127,204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2084 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 250..342
FEATURE:
NAME/KEY: CDS
LOCATION: 250..1791
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 343..1791
US-08-687-399-3

Query Match 84.1%; Score 155.6; DB 2; Length 2084;
Best Local Similarity 97.5%; Pred. No. 2.3e-34;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 18 ATCGATTGTTTGAGAAAGAGAGACCATATAACCTTGTCTGTCATCAGACAGGCT 77
Db 73 ATCGATTGTTTGAGAAAGAGAGACCATATAACCTTGTCTGTCATCAGACAGGCT 132
Qy 78 ATTTTATGCTGTCAGACTGTCGCTGTGTAATAAGGAAAGGAAATGACA 137
Db 133 ATTTTATGCTGTCAGACTGTCGCTGTGTAATAAGGAAATGACA 192
Qy 138 TTATTTTACTGATGATATATATATTTGTAAGAAATG 179
Db 193 TTATTTTACTGATGATATATATTTGTAAGAAATG 234

Search completed: March 19, 2004, 03:17:58
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 00:21:10 ; Search time 215 Seconds
(without alignments)
3655.426 Million cell updates/sec

Title: US-09-834-271A-26

Perfect score: 185
Sequence: 1 ggccttaaggcctgaatc.....tgtataagaaatgagctc 185

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseqn_29Van04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002s.*
- 7: Geneseqn2003as.*
- 8: Geneseqn2003bs.*
- 9: Geneseqn2003cs.*
- 10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	185	100.0	185	2	Aaz23323 B. amylol
2	185	100.0	185	4	Aaf62646 Consensus
3	185	100.0	185	4	Aad09923 Bacillus
4	183.4	99.1	185	2	Aaz23324 B. amylol
5	183.4	99.1	185	4	Aaf62647 Consensus
6	183.4	99.1	185	4	Aad09924 Bacillus
7	180.2	97.4	185	2	Aaz23325 Bacillus
8	180.2	97.4	185	4	Aad09931 Bacillus
9	178.6	96.5	185	2	Aaz23322 B. thurin
10	178.6	96.5	185	4	Aaf62645 Wild type
11	178.6	96.5	185	4	Aad09922 Bacillus
12	159.6	86.3	2166	6	Aad34433 C-termina
13	159.6	86.3	2166	6	Ab140498 DNA const
14	159.6	86.3	2267	6	Aad34434 Subtilisi
15	159.6	86.3	2267	6	Ab140499 DNA const
16	159.6	86.3	2588	6	Aad34436 PamyL ATG
17	159.6	86.3	2588	6	Ab140501 DNA const
18	156.4	84.5	10216	2	Aat39279 Transposo
19	155.6	84.1	162	2	Aat85631 BAN Promo
20	155.6	84.1	249	3	Aaa3677 Bacillus
21	155.6	84.1	249	7	ACA61514 Modified
22	155.6	84.1	270	3	AAA93678 Modified
23	155.6	84.1	537	2	AAQ22579 Sequence

24	155.6	84.1	576	1	AA220046
25	155.6	84.1	2084	2	AAQ88067 Bacillus
26	155.6	84.1	2084	2	AAQ95032 Bacillus
27	155.6	84.1	2084	2	AAQ95032 Bacillus
28	155.6	84.1	2084	2	AAQ95032 Bacillus
29	155.6	84.1	2084	2	AAQ95032 Bacillus
30	155.6	84.1	2084	2	AAQ95032 Bacillus
31	155.6	84.1	2084	2	AAQ95032 Bacillus
32	155.6	84.1	2084	2	AAQ95032 Bacillus
33	155.6	84.1	2084	2	AAQ95032 Bacillus
34	155.6	84.1	2084	2	AAQ95032 Bacillus
35	155.6	84.1	2084	2	AAQ95032 Bacillus
36	155.6	84.1	2084	2	AAQ95032 Bacillus
37	155.6	84.1	2084	2	AAQ95032 Bacillus
38	155.6	84.1	2084	2	AAQ95032 Bacillus
39	155.6	84.1	2084	2	AAQ95032 Bacillus
40	155.6	84.1	2084	2	AAQ95032 Bacillus
41	155.6	84.1	2084	2	AAQ95032 Bacillus
42	155.6	84.1	2084	2	AAQ95032 Bacillus
43	155.6	84.1	2084	2	AAQ95032 Bacillus
44	155.6	84.1	2084	2	AAQ95032 Bacillus
45	155.6	84.1	2084	2	AAQ95032 Bacillus

ALIGNMENTS

RESULT 1
AAZ23323
ID AAZ23323 standard; DNA; 185 BP.
XX
AC AAZ23323;
DT 06-DEC-1999. (first entry)
XX
DE B. amylolifaciens amyQ promoter DNA #1.
XX
KW Tandem promoter; consensus promoter; enzyme production; hormone; amyQ;
KW antibody; reporter; marker gene; cell regulation; alpha-amyase; ss.
XX
OS Bacillus amylolifaciens.
XX
PN WO9943835-A2.
XX
PD 02-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US004360.
XX
PR 26-FEB-1999; 99US-00031442.
XX

(NOVO) NOVO NORDISK BIOTECH INC.
Widner W, Sloma A, Thomas MD;
WPI; 1999-561370/47.
Production of polypeptide in Bacillus using specific promoters,
particularly for producing enzymes.
Claim 44; Page 90; 90pp; English.

This invention describes a novel method for the production of a polypeptide in Bacillus using specific tandem or consensus promoters. The method is used to produce homologous or particularly heterologous proteins, particularly enzymes (specifically serine protease, maltogenic alpha-amyase and pullulanase), but also hormones, antibodies, reporters etc. The specified promoters provide increased expression of the sequence which encodes the polypeptide of the invention. After incorporation of the nucleic acid construct of the invention, any marker gene may be deleted, resulting in a cell that is preferred for environmental and regulatory regions. This sequence represents a Bacillus amylolifaciens alpha-amyase amyQ promoter which is used to describe the method of the invention

applicant

CC region and TATAAT for the -10 region. The method further comprises
 CC isolating the polypeptide from the cultivation medium. The method is
 CC useful for producing a polypeptide in a *Bacillus* strain. The present
 CC sequence is *Bacillus amyloliquefaciens* alpha-amylase (amyQ) consensus
 CC mutated promoter, used in the exemplification of the invention
 XX
 SQ Sequence 185 BP; 63 A; 25 C; 42 G; 55 T; 0 U; 0 Other;

Query Match 100.0%; Score 185; DB 4; Length 185;
 Best Local Similarity 100.0%; Pred. No. 1.6e-41;
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTGT 60
 DB 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTGT 60
 QY 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGA 120
 DB 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGA 120
 QY 121 ATAAAGGGGGTTGACATTATTTTACTGATATGTAATAATATTTGTATAAGAAAATGG 180
 DB 121 ATAAAGGGGGTTGACATTATTTTACTGATATGTAATAATATTTGTATAAGAAAATGG 180
 QY 181 AGCTC 185
 DB 181 AGCTC 185

RESULT 4
 AAZ23324
 ID AAZ23324 standard; DNA; 185 BP.
 XX
 AC AAZ23324;
 XX
 DT 06-DEC-1999 (first entry)
 XX
 DE B. *amyloliquefaciens* amyQ promoter DNA #2.
 XX
 DE Tandem promoter; consensus promoter; enzyme production; hormone; amyQ;
 KW antibody; reporter; marker gene; cell regulation; alpha-amylase; ss.
 XX
 OS *Bacillus amyloliquefaciens*.
 XX
 PN WO943835-A2.
 XX
 PD 02-SEP-1999.
 XX
 PF 26-FEB-1999; 99WO-US004360.
 XX
 PR 26-FEB-1998; 98US-00031442.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 XX
 PI Widner W, Sloma A, Thomas MD;
 XX
 DR WPI; 1998-561370/47.
 XX
 PT Production of polypeptide in *Bacillus* using specific promoters,
 PT particularly for producing enzymes.
 XX
 PS Claim 44; Page 90; 90pp; English.
 XX

This invention describes a novel method for the production of a
 CC polypeptide in *Bacillus* using specific tandem or consensus promoters. The
 CC method is used to produce homologous or particularly heterologous
 CC proteins, particularly enzymes (specifically serine protease, maltogenic
 CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
 CC etc. The specified promoters provide increased expression of the sequence
 CC which encodes the polypeptide of the invention. After incorporation of
 CC the nucleic acid construct of the invention, any marker gene may be
 CC deleted, resulting in a cell that is preferred for environmental and
 CC regulatory regions. This sequence represents a *Bacillus amyloliquefaciens*

CC alpha-amylase amyQ promoter which is used to describe the method of the
 CC invention
 XX
 SQ Sequence 185 BP; 62 A; 25 C; 42 G; 56 T; 0 U; 0 Other;
 Query Match 99.1%; Score 183.4; DB 2; Length 185;
 Best Local Similarity 99.5%; Pred. No. 4.4e-41;
 Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTGT 60
 DB 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAAGAGAACCAATAAAATACCTTGT 60
 QY 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGA 120
 DB 61 CTGTCAATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAATAAAGGA 120
 QY 121 ATAAAGGGGGTTGACATTATTTTACTGATATGTAATAATATTTGTATAAGAAAATGG 180
 DB 121 ATAAAGGGGGTTGACATTATTTTACTGATATGTAATAATATTTGTATAAGAAAATGG 180
 QY 181 AGCTC 185
 DB 181 AGCTC 185

RESULT 5
 AAF62647
 ID AAF62647 standard; DNA; 185 BP.
 XX
 AC AAF62647;
 XX
 DT 03-MAY-2001 (first entry)
 XX
 DE Consensus amyQ promoter #2.
 XX
 KW Pectin acetylase; degrade; plant cell wall; ss.
 XX
 OS Unidentified.
 XX
 PN US6184028-B1.
 XX
 PD 06-FEB-2001.
 XX
 PF 26-AUG-1999; 99US-00384305.
 XX
 PR 26-AUG-1999; 99US-00384305.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 XX
 PI Thomas MD, Brown KW;
 XX
 DR WPI; 2001-190946/19.
 XX

Novel isolated polypeptide having pectin acetylase activity useful
 PT for degrading pectic substances and in degradation or modification of
 PT acetylated pectins and plant cell walls.

Example 9; Col 43-44; 35pp; English.
 PS
 XX The present invention relates to *Bacillus subtilis* pectin acetylase
 CC protein. The invention is useful for degrading a pectic substance. It is
 CC also useful for degrading soluble and insoluble pectins with varying
 CC degrees of esterification, clarification etc. The protein may be used
 CC alone or in combination with other enzymes for the degradation or
 CC modification of acetylated pectins, degradation or modification of plant
 CC cell walls
 XX

SQ Sequence 185 BP; 62 A; 25 C; 42 G; 56 T; 0 U; 0 Other;
 Query Match 99.1%; Score 183.4; DB 4; Length 185;
 Best Local Similarity 99.5%; Pred. No. 4.4e-41;
 Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACACATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACACATATAAAATACCTTGT 60
QY 61 CTGTCAATCAGACAGAGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAAAGGA 120
DB 61 CTGTCAATCAGACAGAGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAAAGGA 120
QY 121 ATAAAGGGGGTGCATATTTTACTGATATGATATATAATTTGTATAGAAATGG 180
DB 121 ATAAAGGGGGTGCATATTTTACTGATATGATATATAATTTGTATAGAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 6

AAD09924
ID AAD09924 standard; DNA; 185 BP.
AC AAD09924;
DT 12-SEP-2001 (first entry)
DE Bacillus amyloliquefaciens alpha-amylase (amyQ) consensus promoter #2.
XX
XX Bacillus cell; tandem promoter; consensus promoter; mutant;
KW polypeptide production; amyL promoter; amyQ promoter; aprH promoter;
KW cryIIIA promoter; subtilisin Carlsberg gene promoter; alpha-amylase; ds.
XX
OS Bacillus amyloliquefaciens.
OS Synthetic.

FH Key Location/Qualifiers

FT mutation replace(135, T)
FT /*tag= a
FT mutation replace(136, T)
FT /*tag= b
FT mutation replace(156, A)
FT /*tag= c
FT /note= "In column 26, a T to A mutation was said to occur
FT at position 116 but no such mutation was found in the
FT sequence shown in the sequence listing of the
FT specification"

PN US6255076-B1.

XX 03-JUL-2001.

XX 26-FEB-1999; 99US-00258377.

XX 26-FEB-1998; 98US-00031442.

XX (NOVO) NOVOZYMES BIOTECH INC.

XX Widner W, Sloma A, Thomas MD;

XX WPI; 2001-440518/47.

XX Producing a polypeptide in a Bacillus strain comprises cultivating a
XX Bacillus cell, which contains a nucleic acid construct comprising a
XX tandem promoter and consensus promoters.

XX Example 18; Col 43-44; 54pp; English.

XX The present invention relates to a method for producing a polypeptide,
XX comprises cultivating a Bacillus cell, which contains a nucleic acid
XX construct comprising a tandem promoter and consensus promoters. The
XX Bacillus is cultivated in a medium conducive for the production of the
XX polypeptide. Each promoter sequence of the tandem promoter is operably
XX linked to a nucleic acid sequence encoding the polypeptide. The tandem

CC promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIIA
CC promoter or subtilisin Carlsberg gene promoter. The consensus promoters
CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter or
CC subtilisin Carlsberg gene promoter have the sequence TTGACA for the -35
CC region and TATAAT for the -10 region. The method further comprises
CC isolating the polypeptide from the cultivation medium. The method is
CC useful for producing a polypeptide in a Bacillus strain. The present
CC sequence is Bacillus amyloliquefaciens alpha-amylase (amyQ) consensus
CC mutated promoter, used in the exemplification of the invention
XX

SQ Sequence 185 BP; 62 A; 25 C; 42 G; 56 T; 0 U; 0 Other;

Query Match 99.1%; Score 183.4; DB 4; Length 185;

Best Local Similarity 99.5%; Pred. NO. 4.4e-41;

Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACACATATAAAATACCTTGT 60
DB 1 GGCCTTAAGGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACACATATAAAATACCTTGT 60
QY 61 CTGTCAATCAGACAGAGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAAAGGA 120
DB 61 CTGTCAATCAGACAGAGGTATTTTATGCTGTCAGACTGTCGGCTGTGTAATAAAAGGA 120
QY 121 ATAAAGGGGGTGCATATTTTACTGATATGATATATAATTTGTATAGAAATGG 180
DB 121 ATAAAGGGGGTGCATATTTTACTGATATGATATATAATTTGTATAGAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 7

AAZ23325

ID AAZ23325 standard; DNA; 185 BP.

XX AC AAZ23325;

XX DT 06-DEC-1999 (first entry)

XX DE Bacillus sp. amyQ promoter consensus DNA.

XX Tandem promoter; consensus promoter; enzyme production; hormone; amyQ;

XX antibody; reporter; marker gene; cell regulation; alpha-amylase; ss.

XX OS Bacillus sp.

XX PN WO9943835-A2.

XX PD 02-SEP-1999.

XX PF 26-FEB-1999; 99WO-US004360.

XX PR 26-FEB-1998; 98US-00031442.

XX PA (NOVO) NOVO NORDISK BIOTECH INC.

XX PI Widner W, Sloma A, Thomas MD;

XX PS WPI; 1999-561370/47.

XX Production of polypeptide in Bacillus using specific promoters,
XX particularly for producing enzymes.

XX Disclosure; Fig 21; 90pp; English.

XX This invention describes a novel method for the production of a
XX polypeptide in Bacillus using specific tandem or consensus promoters. The
XX method is used to produce homologous or particularly heterologous
XX proteins, particularly enzymes (specifically serine protease, maltogenic
XX alpha-amylase and pullulanase), but also hormones, antibodies, reporters
XX etc. The specified promoters provide increased expression of the sequence

CC which encodes the polypeptide of the invention. After incorporation of
 CC the nucleic acid construct of the invention, any marker gene may be
 CC deleted, resulting in a cell that is preferred for environmental and
 CC regulatory regions. This sequence represents the consensus sequence of a
 CC Bacillus sp. alpha-amylase amyQ promoter which is used to describe the
 CC method of the invention
 XX
 SQ Sequence 185 BP; 63 A; 24 C; 42 G; 56 T; 0 U; 0 Other;
 Query Match 97.4%; Score 180.2; DB 2; Length 185;
 Best Local Similarity 98.4%; Pred. No. 3.3e-40;
 Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGT 60
 Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGT 60
 QY 61 CTGTCTATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAATAAAGGA 120
 Db 61 CTGTCTATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAATAAAGGA 120
 QY 121 ATAAAGGGGGTTCACATTTATTTTACTGATATGTAATAATATTTGTAAGAAATGG 180
 Db 121 ATAAAGGGGGTTCATTTATTTTACTGATATGTAATAATATTTGTAAGAAATGG 180
 QY 181 AGCTC 185
 Db 181 AGCTC 185
 RESULT 8
 AAD09931
 ID AAD09931 standard; DNA; 185 BP.
 AC AAD09931;
 DT 12-SEP-2001 (first entry)
 XX Bacillus amyloliquefaciens amyQ promoter, alternative version.
 DE
 XX Bacillus cell; tandem promoter; consensus promoter;
 KW polypeptide production; amyL promoter; amyQ promoter; aprH promoter;
 KW cryIIIA promoter; subtilisin Carlsberg gene promoter; alpha-amylase; ds.
 XX
 OS Bacillus amyloliquefaciens.
 XX
 FH Key Location/Qualifiers
 FT misc_feature 116
 FT /*tag= a
 FT /note= "Base T is found at this location in the sequence
 FT shown in the sequence listing of the specification
 FT (AAD09922)."
 XX
 PN US6255076-B1.
 XX
 PD 03-JUL-2001.
 XX
 PF 26-FEB-1999; 99US-00258377.
 XX
 PR 26-FEB-1998; 98US-00031442.
 XX
 PA (NOVO) NOVOZYMES BIOTECH INC.
 XX
 PI Widner W, Sloma A, Thomas MD;
 XX
 DR WPI; 2001-440518/47.
 XX
 XX Producing a polypeptide in a Bacillus strain comprises cultivating a
 PT Bacillus cell, which contains a nucleic acid construct comprising a
 PT tandem promoter and consensus promoters.
 XX
 XX Example 18; Fig 21; 54pp; English.
 PS
 XX

CC The present invention relates to a method for producing a polypeptide,
 CC comprises cultivating a Bacillus cell, which contains a nucleic acid
 CC construct comprising a tandem promoter and consensus promoters. The
 CC Bacillus is cultivated in a medium conducive for the production of the
 CC polypeptide. Each promoter sequence of the tandem promoter is operably
 CC linked to a nucleic acid sequence encoding the polypeptide. The tandem
 CC promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIIA
 CC promoter or subtilisin Carlsberg gene promoter. The consensus promoters
 CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter or
 CC subtilisin Carlsberg gene promoter have the sequence TTGACA for the -35
 CC region and TATAAT for the -10 region. The method further comprises
 CC isolating the polypeptide from the cultivation medium. The method is
 CC useful for producing a polypeptide in a Bacillus strain. The present
 CC sequence is an alternative of Bacillus amyloliquefaciens alpha-amylase
 CC (amyQ) promoter, used in the exemplification of the invention. Note: This
 CC sequence is stated as being the same as that shown as SEQ ID NO 25
 CC (AAD09922) in column 43-44 of the specification. However the sequences
 CC differ by at position 116
 XX
 SQ Sequence 185 BP; 63 A; 24 C; 42 G; 56 T; 0 U; 0 Other;
 Query Match 97.4%; Score 180.2; DB 4; Length 185;
 Best Local Similarity 98.4%; Pred. No. 3.3e-40;
 Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGT 60
 Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAATACCTTGT 60
 QY 61 CTGTCTATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAATAAAGGA 120
 Db 61 CTGTCTATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAATAAAGGA 120
 QY 121 ATAAAGGGGGTTCACATTTATTTTACTGATATGTAATAATATTTGTAAGAAATGG 180
 Db 121 ATAAAGGGGGTTCATTTATTTTACTGATATGTAATAATATTTGTAAGAAATGG 180
 QY 181 AGCTC 185
 Db 181 AGCTC 185
 RESULT 9
 AAZ23322
 ID AAZ23322 standard; DNA; 185 BP.
 AC AAZ23322;
 DT 06-DEC-1999 (first entry)
 DE B. thuringiensis wild-type amyQ promoter DNA.
 XX
 XX Tandem promoter; consensus promoter; enzyme production; hormone; amyQ;
 KW antibody; reporter; marker gene; cell regulation; alpha-amylase; ss.
 XX
 OS Bacillus thuringiensis.
 XX
 PN WO9943835-A2.
 XX
 PD 02-SEP-1999.
 XX
 PF 26-FEB-1999; 99WC-US004360.
 XX
 PR 26-FEB-1998; 98US-00031442.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 XX
 PI Widner W, Sloma A, Thomas MD;
 XX
 DR WPI; 1999-561370/47.
 XX
 XX Production of polypeptide in Bacillus using specific promoters,
 PT particularly for producing enzymes.

XX Disclosure; Page 90; 90pp; English.

XX This invention describes a novel method for the production of a

CC polypeptide in *Bacillus* using specific tandem or consensus promoters. The

CC method is used to produce homologous or particularly heterologous

CC proteins, particularly enzymes (specifically serine protease, maltogenic

CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters

CC etc. The specified promoters provide increased expression of the sequence

CC which encodes the polypeptide of the invention. After incorporation of

CC the nucleic acid construct of the invention, any marker gene may be

CC deleted, resulting in a cell that is preferred for environmental and

CC regulatory regions. This sequence represents a *Bacillus thuringiensis*

CC alpha-amylase amyQ promoter which is used to describe the method of the

XX invention

XX Sequence 185 BP; 62 A; 24 C; 42 G; 57 T; 0 U; 0 Other;

Query Match 96.5%; Score 178.6; DB 2; Length 185;

Best Local Similarity 97.8%; Pred. No. 9.2e-40;

Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGT 60

DB 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGT 60

QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAGAAAGGA 120

DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAGAAATAGGA 120

QY 121 ATAAAGGGGGTTCACATTATTTTACTGATATGTAATAATATTTGTATAGAAAATGG 180

DB 121 ATAAAGGGGGTTCATTATTTTACTGATATGTAATAATATTTGTATAGAAAATGG 180

QY 181 AGCTC 185

DB 181 AGCTC 185

RESULT 10

AAF62645

ID AAF62645 standard; DNA; 185 BP.

XX AAF62645;

XX 03-MAY-2001 (first entry)

XX Wild type amyQ promoter.

XX Pectin acetyltransferase; degrade; plant cell wall; ss.

XX Unidentified.

XX US6184028-B1.

XX 06-FEB-2001.

XX 26-AUG-1999; 99US-00384305.

XX 26-AUG-1999; 99US-00384305.

XX (NOVO) NOVO NORDISK BIOTECH INC.

XX Thomas MD, Brown KM;

XX WPI; 2001-190946/19.

XX Novel isolated polypeptide having pectin acetyltransferase activity useful

PT for degrading pectic substances and in degradation or modification of

PT acetylated pectins and plant cell walls.

XX Example 9; Fig 9; 35pp; English.

CC The present invention relates to *Bacillus subtilis* pectin acetyltransferase

CC protein. The invention is useful for degrading a pectic substance. It is

CC also useful for degrading soluble and insoluble pectins with varying

CC degrees of esterification, clarification etc. The protein may be used

CC alone or in combination with other enzymes for the degradation or

CC modification of acetylated pectins, degradation or modification of plant

CC cell walls

XX Sequence 185 BP; 62 A; 24 C; 42 G; 57 T; 0 U; 0 Other;

Query Match 96.5%; Score 178.6; DB 4; Length 185;

Best Local Similarity 97.8%; Pred. No. 9.2e-40;

Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGT 60

DB 1 GGCCTTAAGGGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGT 60

QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAGAAAGGA 120

DB 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAGAAATAGGA 120

QY 121 ATAAAGGGGGTTCACATTATTTTACTGATATGTAATAATATTTGTATAGAAAATGG 180

DB 121 ATAAAGGGGGTTCATTATTTTACTGATATGTAATAATATTTGTATAGAAAATGG 180

QY 181 AGCTC 185

DB 181 AGCTC 185

RESULT 11

AAD09922

ID AAD09922 standard; DNA; 185 BP.

XX AAD09922;

XX 12-SEP-2001 (first entry)

XX *Bacillus amyloliquefaciens* alpha-amylase (amyQ) promoter.

XX *Bacillus* cell; tandem promoter; consensus promoter;

XX polypeptide production; amyL promoter; amyQ promoter; aprH promoter;

XX cryIIIA promoter; subtilisin Carlsberg gene promoter; alpha-amylase; ds.

XX *Bacillus amyloliquefaciens*.

XX US6255076-B1.

XX 03-JUL-2001.

XX 26-FEB-1999; 99US-00258377.

XX 26-FEB-1998; 98US-00031442.

XX (NOVO) NOVOZYMES BIOTECH INC.

XX Widner W, Sloma A, Thomas MD;

XX WPI; 2001-440518/47.

XX Producing a polypeptide in a *Bacillus* strain comprises cultivating a

PT *Bacillus* cell, which contains a nucleic acid construct comprising a

PT tandem promoter and consensus promoters.

XX Example 18; Col 43-44; 54pp; English.

XX The present invention relates to a method for producing a polypeptide,

CC comprises cultivating a *Bacillus* cell, which contains a nucleic acid

CC construct comprising a tandem promoter and consensus promoters. The

CC *Bacillus* is cultivated in a medium conducive for the production of the

CC polypeptide. Each promoter sequence of the tandem promoter is operably

CC linked to a nucleic acid sequence encoding the polypeptide. The tandem

CC promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIIA
 CC promoter or subtilisin Carlsberg gene promoter. The consensus promoters
 CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter or
 CC subtilisin Carlsberg gene promoter have the sequence TTGACA for the -35
 CC region and TATAAT for the -10 region. The method further comprises
 CC isolating the polypeptide from the cultivation medium. The method is
 CC useful for producing a polypeptide in a Bacillus strain. The present
 CC sequence is Bacillus anoliquefaciens alpha-amylase (amyQ) promoter,
 CC used in the exemplification of the invention. Note: This sequence is
 CC stated as being the same as that shown as SEQ ID NO 25 (AAB09331) in
 CC figure 21 of the specification. However the sequences differ by at
 CC position 116
 XX
 XX Sequence 185 BP; 62 A; 24 C; 42 G; 57 T; 0 U; 0 Other;

Query Match 96.5%; Score 178.6; DB 4; Length 185;
 Best Local Similarity 97.8%; Pred. No. 9.2e-40;
 Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GGCCTTAAGGCCCTGCATCGATTGTTTGAGAAAGAGACCATATAAAATACCTTGT 60
 DB 1 GGCCTTAAGGCCCTGCATCGATTGTTTGAGAAAGAGACCATATAAAATACCTTGT 60
 QY 61 CTGTCATCAGACAGGGGTATTTTATGCTGTCAGACTGTCGCGTGTAAAAAAGGA 120
 DB 61 CTGTCATCAGACAGGGGTATTTTATGCTGTCAGACTGTCGCGTGTAAAAAATAGGA 120
 QY 121 ATAAAGGGGGTTCACATTTATTTTACTGATGATGATATATATTTGTAAGAAATAGG 180
 DB 121 ATAAAGGGGGTTCATTTATTTTACTGATGATGATGATATATTTGTAAGAAATAGG 180
 QY 181 AGCTC 185
 DB 181 AGCTC 185

RESULT 12
 AAD34433
 ID AAD34433 standard; DNA; 2166 BP.
 XX
 XX AAD34433;
 DT 29-AUG-2003 (revised)
 DT 16-JUL-2002 (first entry)
 DE C-terminal Subtilisin 309-CI-2A fusion protein encoding construct A DNA.
 KW Protease; screening; barley; chymotrypsin inhibitor; CI-2A; laundry;
 KW detergent industry; dishwash; fabric softener; subtilisin 309; savinase;
 KW chimeric; ds.
 XX
 XX Hordeum sp.
 OS Bacillus lentus.
 OS Chimeric.
 XX
 XX WO200218588-A1.
 PN
 XX
 XX 07-MAR-2002.
 XX
 XX 17-JUL-2001; 2001WO-DK000503.
 XX
 XX 29-AUG-2000; 2000DK-00001273.
 PR
 XX 05-SEP-2000; 2000US-0230017P.
 XX
 XX (NOVO) NOVOZYMES AS.
 PA
 XX Pedersen PE, Norregaard-Madsen M;
 PI
 XX WPI; 2002-329776/36.
 DR
 XX Screening protease/protease inhibitor gene library for gene encoding the
 PT same, by constructing host cell having the gene, cultivating cell
 PT dissociating complex formed between them, selecting protein and isolating

PT gene.
 XX Example 1; Page 47-48; 54pp; English.
 CC The invention relates to a method for screening a protease/protease
 CC inhibitor gene library for a gene encoding a protease of interest. The
 CC method comprises introducing a first gene of the protease gene library
 CC and a second gene encoding a protease inhibitor into a host cell;
 CC culturing the host cell, wherein the cell expresses the first and the
 CC second genes to produce a complex of a protease and the inhibitor;
 CC dissociating the inhibitor from the complex and selecting the protease of
 CC interest and isolating the encoding gene. A complex may be used for a
 CC number of industrial applications, in particular within the detergent
 CC industry, thus, relates to a cleaning or detergent composition,
 CC preferably a laundry or dishwash composition comprising the complex,
 CC where hand or machine laundry detergent composition including a laundry
 CC additive composition suitable for pre-treatment of stained fabrics and a
 CC rinse added composition, or be formulated as a detergent composition for
 CC use in general household hard surface cleaning operations or be
 CC formulated for hand or machine dish washing operations. The present
 CC sequence is a construct encoding a fusion protein containing C-terminal
 CC region of subtilisin 309 (savinase) and barley chymotrypsin inhibitor CI-
 CC 2A protein used in the invention. (Updated on 29-AUG-2003 to standardise
 CC OS field)
 XX
 XX Sequence 2166 BP; 610 A; 442 C; 582 G; 532 T; 0 U; 0 Other;

Query Match 86.3%; Score 159.6; DB 6; Length 2166;
 Best Local Similarity 97.6%; Pred. No. 2.8e-34;
 Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 14 TCGAATCGATTGTTTGAGAAAGAGACCATATAAAATACCTTGTCTGTCATCAGACA 73
 DB 172 TCGAATCGATTGTTTGAGAAAGAGACCATATAAAATACCTTGTCTGTCATCAGACA 231
 QY 74 GGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAAAAAGGAATAAAGGGGGTT 133
 DB 232 GGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAAAAAGGAATAAAGGGGGTT 291
 QY 134 GACATTTTACTGATGATGATATATATTTGTAAGAAATG 179
 DB 292 GTTATTTTACTGATGATGATATATATTTGTAAGAAATG 337

RESULT 13
 ABL40498
 ID ABL40498 standard; DNA; 2166 BP.
 XX
 XX ABL40498;
 XX
 XX 17-JUN-2002 (first entry)
 DT
 XX DNA construct A comprising Savinase-CI-2A fusion sequence.
 DE
 XX Protease; allergenicity; detergent; subtilase; S11; S12; additive;
 KW protease inhibitor; barley; chymotrypsin; CI-2A; fusion protein;
 KW subtilisin; savinase; apr gene; ds.
 XX
 XX Synthetic.
 OS Bacillus sp.
 OS Hordeum sp.
 XX
 XX WO200216619-A1.
 PN
 XX
 XX 28-FEB-2002.
 PD
 XX 09-JUL-2001; 2001WO-DK000479.
 PF
 XX 21-AUG-2000; 2000DK-00001233.
 PR
 XX 24-AUG-2000; 2000US-0227501P.
 XX
 XX (NOVO) NOVOZYMES AS.
 PA
 XX

PI Pedersen PE;
 XX WPI; 2002-242031/29.
 XX Producing protease-inhibitor compositions with reduced allergenicity,
 PT stability and activity for use in detergent compositions.
 XX
 PS Example 1; Page 45-46; 52pp; English.
 CC The invention provides a method for producing a protease-inhibitor
 CC complex with reduced allergenicity for use in detergents. The method
 CC involves (a) constructing a fusion polynucleotide sequence in frame, the
 CC sequence comprising a first gene encoding a protease (preferably a
 CC subtilase S11 or S12) and a second gene encoding a protease inhibitor;
 CC (b) introducing the sequence into a host cell; and (c) cultivating the
 CC host cell (the host cell expresses the sequence and produces a non-
 CC covalently linked complex of the protease and the inhibitor). The
 CC protease-inhibitor complex produced by the method may be used as an
 CC nucleotide sequence of a DNA construct comprising the savinase-CI-2A
 CC fusion sequence
 XX
 SQ Sequence 2166 BP; 610 A; 442 C; 582 G; 532 T; 0 U; 0 Other;
 Query Match 86.3%; Score 159.6; DB 6; Length 2166;
 Best Local Similarity 97.6%; P-red. No. 2.8e-34;
 Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 14 TGCATCGATTGTTGAGAAAGAGAGACACCAATAAATACCTGTCTGTCATCAGACA 73
 DB 172 TGCATCGATTGTTGAGAAAGAGAGACACCAATAAATACCTGTCTGTCATCAGACA 231
 QY 74 GGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAAAAAGGAATAAGGGGGTT 133
 DB 232 GGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAAAAAGGAATAAGGGGGTT 291
 QY 134 GACATTTATTTTACTGATGATATATATATATATATATATATATATATATATAT 179
 DB 292 GTTATTTATTTTACTGATGATATATATATATATATATATATATATATATATAT 337
 RESULT 14
 AAD34434
 ID AAD34434 standard; DNA; 2267 BP.
 AC AAD34434;
 XX
 XX 29-AUG-2003 (revised)
 DT 16-JUL-2002 (first entry)
 XX Subtilisin 309-CI-2A transcriptional product encoding construct B DNA.
 XX
 XX Protease; screening; barley; chymotrypsin inhibitor; CI-2A; laundry;
 KW detergent industry; dishwash; fabric softener; subtilisin 309; savinase;
 KW chimeric; ds.
 XX
 XX Hordeum sp.
 OS Bacillus lentus.
 OS Chimeric.
 OS
 PN WO200218588-A1.
 XX
 XX 07-MAR-2002.
 XX
 XX 17-JUL-2001; 2001WO-DK000503.
 PF
 XX 29-AUG-2000; 2000DK-00001273.
 PR 05-SEP-2000; 2000US-0230017P.
 XX
 XX (NOVO) NOVOZYMES AS.
 PA
 XX Pedersen PE, Norregaard-Madsen M;
 PI WPI; 2002-329776/36.
 DR

XX Screening protease/protease inhibitor gene library for gene encoding the
 PT same, by constructing host cell having the gene, cultivating cell,
 PT dissociating complex formed between them, selecting protein and isolating
 PT gene.
 XX
 PS Example 1; Page 48; 54pp; English.
 CC The invention relates to a method for screening a protease/protease
 CC inhibitor gene library for a gene encoding a protease of interest. The
 CC method comprises introducing a first gene of the protease gene library
 CC and a second gene encoding a protease inhibitor into a host cell;
 CC cultivating the host cell, wherein the cell expresses the first and the
 CC second genes to produce a complex of a protease and the inhibitor;
 CC dissociating the inhibitor from the complex and selecting the protease of
 CC interest and isolating the encoding gene. A complex may be used for a
 CC number of industrial applications, in particular within the detergent
 CC industry, thus, relates to a cleaning or detergent composition,
 CC preferably a laundry or dishwash detergent composition comprising a laundry
 CC where hand or machine laundry detergent composition including a laundry
 CC additive composition suitable for pre-treatment of stained fabrics and a
 CC rinse added composition, or be formulated as a detergent composition for
 CC use in general household hard surface cleaning operations or be
 CC formulated for hand or machine dish washing operations. The present
 CC sequence is a construct encoding a transcriptional fusion product
 CC containing C-terminal region of subtilisin 309 (savinase) and barley
 CC chymotrypsin inhibitor CI-2A protein used in the invention. (Updated on
 CC 29-AUG-2003 to standardise OS field)
 XX
 SQ Sequence 2267 BP; 640 A; 462 C; 601 G; 564 T; 0 U; 0 Other;
 Query Match 86.3%; Score 159.6; DB 6; Length 2267;
 Best Local Similarity 97.6%; P-red. No. 2.8e-34;
 Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 14 TGCATCGATTGTTGAGAAAGAGAGACACCAATAAATACCTGTCTGTCATCAGACA 73
 DB 172 TGCATCGATTGTTGAGAAAGAGAGACACCAATAAATACCTGTCTGTCATCAGACA 231
 QY 74 GGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAAAAAGGAATAAGGGGGTT 133
 DB 232 GGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAAAAAGGAATAAGGGGGTT 291
 QY 134 GACATTTATTTTACTGATGATATATATATATATATATATATATATATATATAT 179
 DB 292 GTTATTTATTTTACTGATGATATATATATATATATATATATATATATATATAT 337
 RESULT 15
 ABL40499
 ID ABL40499 standard; DNA; 2267 BP.
 AC ABL40499;
 XX
 XX 17-JUN-2002 (first entry)
 DT
 XX DNA construct B comprising Savinase-CI-2A fusion sequence.
 DE
 XX Protease; allergenicity; detergent; subtilase; S11; S12; additive;
 KW protease inhibitor; barley; chymotrypsin; CI-2A; fusion protein;
 KW subtilisin; savinase; apr gene; ds.
 XX
 XX Synthetic.
 OS Bacillus sp.
 OS Hordeum sp.
 OS
 XX WO200216619-A1.
 PN
 XX 28-FEB-2002.
 XX
 XX 09-JUL-2001; 2001WO-DK000479.
 PF
 XX 21-AUG-2000; 2000DK-00001233.
 PR

PR 24-AUG-2000; 2000US-0227501P.
XX (NOVO) NOVOZYMES AS.
PA Pedersen PE;
XX WPI; 2002-242031/29.
XX
XX Producing protease-inhibitor compositions with reduced allergenicity,
PT stability and activity for use in detergent compositions.
XX
XX Example 1; Page 46; 52pp; English.
XX
XX The invention provides a method for producing a protease-inhibitor
XX complex with reduced allergenicity for use in detergents. The method
XX involves (a) constructing a fusion polynucleotide sequence in frame, the
XX sequence comprising a first gene encoding a protease (preferably a
XX subtilase S11 or S12) and a second gene encoding a protease inhibitor;
XX (b) introducing the sequence into a host cell; and (c) cultivating the
XX host cell (the host cell expresses the sequence and produces a non-
XX covalently linked complex of the protease and the inhibitor). The
XX protease-inhibitor complex produced by the method may be used as an
XX nucleotide sequence of a DNA construct comprising the savinase-CI-2A
XX fusion sequence
XX
SQ Sequence 2267 BP; 640 A; 462 C; 601 G; 564 T; 0 U; 0 Other;

Query Match 86.3%; Score 159.6; DB 6; Length 2267;
Best Local Similarity 97.6%; Pred. No. 2.8e-34;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 TGCATCGATGTTTGAGAAAGAGAGAGACCATTAATAATACCTTCTCTCATCAGACA 73
Db 172 TGCATCGATGTTTGAGAAAGAGAGAGACCATTAATAATACCTTCTCTCATCAGACA 231

QY 74 GGGTATTTTATGCTGTCAGACTGCCGCTGTGTAATAAAGGAATAAAGGGGGTT 133
Db 232 GGGTATTTTATGCTGTCAGACTGCCGCTGTGTAATAAAGGAATAAAGGGGGTT 291

QY 134 GACATATTTTACTGATATGATATATAATATTTTGTATAGAAAATG 179
Db 292 GTTATTTTACTGATATGATATATAATATTTTGTATAGAAAATG 337

Search completed: March 19, 2004, 01:38:17
Job time : 218 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 00:23:20 ; Search time 1078.5 Seconds
(without alignments)
7434.827 Million cell updates/sec

Title: US-09-834-271A-26
Perfect score: 185
Sequence: 1 ggccttaaggccgtgcattc.....tgataaagaagtgagctc 185

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.man.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	185	100.0	185	6	ARI60502 Sequence
2	185	100.0	185	6	AX088920 Sequence
3	185	100.0	185	6	BD136798 Process f
4	183.4	99.1	185	6	ARI60503 Sequence
5	183.4	99.1	185	6	AX088921 Sequence
6	183.4	99.1	185	6	BD136799 Process f
7	178.6	96.5	185	6	ARI60501 Sequence
8	178.6	96.5	185	6	AX088919 Sequence
9	178.6	96.5	185	6	BD136797 Process f
10	155.6	84.1	162	6	I40596 Sequence 24
11	155.6	84.1	162	6	I40867 Sequence 24
12	155.6	84.1	162	6	I40895 Sequence 24
13	155.6	84.1	162	6	I56837 Sequence 24
14	155.6	84.1	249	6	BD174625 Modified
15	155.6	84.1	537	1	BAAAYL
16	155.6	84.1	576	6	A00155 Nucleotide
17	155.6	84.1	576	6	A00156 Nucleotide
18	155.6	84.1	576	6	A00157 Nucleotide
19	155.6	84.1	576	6	A00158 Nucleotide
20	155.6	84.1	576	6	A00604 B.amyloliqu
21	155.6	84.1	576	6	A00605 B.amyloliqu
22	155.6	84.1	2084	1	BACAAAM
23	155.6	84.1	2084	6	AR008285 Sequence
24	155.6	84.1	2084	6	AR037275 Sequence
25	155.6	84.1	2084	6	AR052145 Sequence
26	155.6	84.1	2084	6	AR07551 Sequence
27	155.6	84.1	2084	6	AR129916 Sequence
28	155.6	84.1	2084	6	AR137905 Sequence
29	155.6	84.1	2084	6	AR143262 Sequence
30	155.6	84.1	2084	6	BD230131 Alpha-amy
31	155.6	84.1	2084	6	AR224261 Sequence
32	155.6	84.1	2084	6	AR285306 Sequence
33	155.6	84.1	2084	6	AR428861 Sequence
34	155.6	84.1	2084	6	AX244195 Sequence
35	155.6	84.1	2084	6	AX305020 Sequence
36	155.6	84.1	2084	6	AX339268 Sequence
37	155.6	84.1	2084	6	AX370717 Sequence
38	155.6	84.1	2084	6	AX601810 Sequence
39	155.6	84.1	2084	6	BD077082 Alpha-amy
40	155.6	84.1	2084	6	BD085805 Alpha-amy
41	155.6	84.1	2604	6	AR168304 Sequence
42	155.6	84.1	2604	6	BD249244 Alpha-amy
43	155.6	84.1	2604	6	AR215288 Sequence
44	155.6	84.1	2604	6	AX036892 Sequence
45	155.6	84.1	2604	6	BD137366 Enzymatic

ALIGNMENTS

RESULT 1
ARI60502
LOCUS ARI60502 185 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 26 from patent US 6255076.
ACCESSION ARI60502
VERSION ARI60502.1 GI:16224477
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 185)
AUTHORS Widner,W., Sloma,A. and Thomas,M.D.
TITLE Methods for producing a polypeptide in a Bacillus cell
JOURNAL Patent: US 6255076-A 26 03-JUL-2001;
FEATURES Location/Qualifiers

Pred. No. is the number of results predicted by chance to have a

```
source          1. .185
                /organism="unknown"
                /mol_type="unassigned DNA"

ORIGIN
Query Match          100.0%; Score 185; DB 6; Length 185;
Best Local Similarity 100.0%; Pred. No. 7.2e-35;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTGAGAAAAGAGAACCACTAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTGAGAAAAGAGAACCACTAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTTCGACATTTTACTGATATGATATATATATTTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTCGACATTTTACTGATATGATATATATATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 2
AX088920          185 bp      DNA      linear      PAT 17-MAR-2001
LOCUS
DEFINITION
Sequence 21 from Patent WO0114534.
ACCESSION
AX088920
VERSION
AX088920.1 GI:13397679
SOURCE
Bacillus subtilis
ORGANISM
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
1 Thomas, M.D. and Brown, K.M.
Polypeptides having pectin acetyltransferase activity and nucleic
acids encoding same
Patent: WO 0114534-A 21 01-MAR-2001;
Novozymes Biotech, Inc. (US)
FEATURES
source
1. .185
/organism="Bacillus subtilis"
/mol_type="unassigned DNA"
/db_xref="taxon:1423"

ORIGIN
Query Match          100.0%; Score 185; DB 6; Length 185;
Best Local Similarity 100.0%; Pred. No. 7.2e-35;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTGAGAAAAGAGAACCACTAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTGAGAAAAGAGAACCACTAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTTCGACATTTTACTGATATGATATATATATTTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTCGACATTTTACTGATATGATATATATATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 3
BD136798          185 bp      DNA      linear      PAT 18-SEP-2002
LOCUS
```

```
DEFINITION
Process for producing polypeptide in bacillus cells.
ACCESSION
BD136798
VERSION
BD136798.1 GI:23231743
KEYWORDS
JP 2002504379-A/26.
SOURCE
unidentified
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 185)
Widner, W., Sloma, A. and Thomas, M.D.
Process for producing polypeptide in bacillus cells
Patent: JP 2002504379-A 26 12-FEB-2002;
NOVO NORDISK BIOTECH INC
OS
Bacillus
PN JP 2002504379-A/26
PD 12-FEB-2002
PF 26-FEB-1999 JP 2000533574
PR 26-FEB-1998 US 09/031442
PI WILLIAM WIDNER, ALAN SLOMA, MICHAEL D THOMAS
PC C12N15/09, C12N1/21, C12N9/54//C07K14/325, (C12N1/21, C12N1/07),
PC C12N15/00
CC Process for producing polypeptide in bacillus cells FH Key
Location/Qualifiers
FT source 1. .185
/organism="Bacillus".
FEATURES
source
1. .185
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match          100.0%; Score 185; DB 6; Length 185;
Best Local Similarity 100.0%; Pred. No. 7.2e-35;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTGAGAAAAGAGAACCACTAAATACCTTGT 60
DB 1 GGCCTTAAGGCGCTGCAATCGATTGTTGAGAAAAGAGAACCACTAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGA 120
DB 61 CTGTCATCAGACAGGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAAAAAAGGA 120
QY 121 ATAAAGGGGGTTCGACATTTTACTGATATGATATATATATTTTGTATAGAAAATGG 180
DB 121 ATAAAGGGGGTTCGACATTTTACTGATATGATATATATATTTGTATAGAAAATGG 180
QY 181 AGCTC 185
DB 181 AGCTC 185

RESULT 4
AR160503          185 bp      DNA      linear      PAT 17-OCT-2001
LOCUS
DEFINITION
Sequence 27 from patent US 6255076.
ACCESSION
AR160503
VERSION
AR160503.1 GI:16224480
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 185)
Widner, W., Sloma, A. and Thomas, M.D.
Methods for producing a polypeptide in a Bacillus cell
Patent: US 6255076-A 27 03-JUL-2001;
Novozymes Biotech, Inc. (US)
FEATURES
source
1. .185
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match          99.1%; Score 183.4; DB 6; Length 185;
```



```
Best Local Similarity 99.5%; Pred. No. 1.8e-34; Mismatches 0; Indels 0; Gaps 0;
Matches 184; Conservative 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAAAAAGGA 120
Db 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAAAAATAGGA 120
QY 121 ATAAAGGGGGTTCACATTTTACTGATATGTATAATAATTTGTATAAGAAATGG 180
Db 121 ATAAAGGGGGTTCACATTTTACTGATATGTATAATAATTTGTATAAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 5
AX088921
LOCUS AX088921 185 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 22 from Patent WO0114534.
ACCESSION AX088921
VERSION AX088921.1 GI:13397680
KEYWORDS
SOURCE
ORGANISM
Bacillus subtilis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
1 Thomas, M.D. and Brown, K.M.
Polypeptides having pectin acetyltransferase activity and nucleic
acids encoding same
Patent: WO 0114534-A 22 01-MAR-2001;
Novozymes Biotech, Inc. (US)
FEATURES
source
1. 185
/organism="Bacillus subtilis"
/mol_type="unassigned DNA"
/db_xref="taxon:1423"
ORIGIN
Query Match 99.1%; Score 183.4; DB 6; Length 185;
Best Local Similarity 99.5%; Pred. No. 1.8e-34;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAAAAAGGA 120
Db 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAAAAATAGGA 120
QY 121 ATAAAGGGGGTTCACATTTTACTGATATGTATAATAATTTGTATAAGAAATGG 180
Db 121 ATAAAGGGGGTTCACATTTTACTGATATGTATAATAATTTGTATAAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 6
BD136799
LOCUS BD136799 185 bp DNA linear PAT 18-SEP-2002
DEFINITION Process for producing polypeptide in bacillus cells.
ACCESSION BD136799
VERSION BD136799.1 GI:23231744
KEYWORDS
SOURCE
ORGANISM
unidentified
```

```
unclassified.
1 (bases 1 to 185)
Widner, W., Sloma, A. and Thomas, M.D.
AUTHORS
TITLE Process for producing polypeptide in bacillus cells
JOURNAL Patent: JP 2002504379-A 27 12-FEB-2002;
NOVO NORDISK BIOTECH INC
COMMENT
OS Bacillus
PN JP 2002504379-A/27
PD 12-FEB-2002
PF 26-FEB-1999 JP 2000533574
PR 26-FEB-1998 US 09/031442
PI WILLIAM WIDNER, ALAN SLOMA, MICHAEL D THOMAS
PC C12N15/00, C12N1/21, C12N9/54//C07K14/325, (C12N1/21, C12R1:07),
PC C12N15/00
CC Process for producing polypeptide in bacillus cells FH Key
FT source
Location/Qualifiers
1. 185
/organism="Bacillus"
FEATURES
source
1. 185
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 99.1%; Score 183.4; DB 6; Length 185;
Best Local Similarity 99.5%; Pred. No. 1.8e-34;
Matches 184; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAAAAAGGA 120
Db 61 CTGTCATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAAAAATAGGA 120
QY 121 ATAAAGGGGGTTCACATTTTACTGATATGTATAATAATTTGTATAAGAAATGG 180
Db 121 ATAAAGGGGGTTCACATTTTACTGATATGTATAATAATTTGTATAAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 7
ARI60501
LOCUS ARI60501 185 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 25 from patent US 6255076.
ACCESSION ARI60501
VERSION ARI60501.1 GI:16224474
KEYWORDS
SOURCE
ORGANISM
Unknown.
UNCLASSIFIED.
REFERENCE
1 (bases 1 to 185)
Widner, W., Sloma, A. and Thomas, M.D.
AUTHORS
TITLE Methods for producing a polypeptide in a Bacillus cell
JOURNAL Patent: US 6255076-A 25 03-JUL-2001;
FEATURES
source
1. 185
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 96.5%; Score 178.6; DB 6; Length 185;
Best Local Similarity 97.8%; Pred. No. 2.6e-33;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAAGAGAACCACTAAATAATACCTTGT 60
```

```
QY 61 CTGTCATCAGACAGGGTATTTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
Db 61 CTGTCATCAGACAGGGTATTTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
QY 121 ATAAAGGGGGTGCATTTATTTACTGATATGATATATAATTTGATATAGAAATGG 180
Db 121 ATAAAGGGGGTGTATTATTTACTGATATGATATATAATTTGATATAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 8
AX088919
LOCUS 185 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 20 from Patent WO0114534.
ACCESSION AX088919
VERSION AX088919.1 GI:13397678
KEYWORDS Bacillus subtilis
SOURCE Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
ORGANISM
REFERENCE 1
AUTHORS Thomas, M. D. and Brown, K. M.
TITLE Polypeptides having pectin acetyltransferase activity and nucleic acids encoding same
JOURNAL Patent: WO 0114534-A 20 01-MAR-2001;
Novozymes Biotech, Inc. (US)
FEATURES
Source
1. .185
/organism="Bacillus subtilis"
/mol_type="unassigned DNA"
/db_xref="taxon:1423"

ORIGIN
Query Match 96.5%; Score 178.6; DB 6; Length 185;
Best Local Similarity 97.8%; Pred. No. 2.6e-33;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGT 60
Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
Db 61 CTGTCATCAGACAGGGTATTTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
QY 121 ATAAAGGGGGTGCATTTATTTACTGATATGATATATAATTTGATATAGAAATGG 180
Db 121 ATAAAGGGGGTGTATTATTTACTGATATGATATATAATTTGATATAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 10
LOCUS 162 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 24 from patent US 5621089.
ACCESSION I40596
VERSION I40596.1 GI:2082888
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 162)
AUTHORS Sloma, A. P.; Outtrup, H.; Dammann, C. and Aaslyng, D. A.
TITLE Nucleic acid constructs for the production of a Bacillus alkaline protease
JOURNAL Patent: US 5621089-A 24 15-APR-1997;
FEATURES
Source
1. .162
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 84.1%; Score 155.6; DB 6; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.1e-27;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGTGTATCAGACAGGT 77
Db 1 ATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGTGTATCAGACAGGT 60
QY 78 ATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGAATAGGGGGTTCACA 137
Db 61 ATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGAATAGGGGGTTCGTA 120
```

```
OS Bacillus
PN JP 2002504379-A/25
PD 12-FEB-2002
PF 26-FEB-1999 JP 2000533574
PR 26-FEB-1998 US 09/031442
PI WILLIAM WIDNER, ALAN SLOMA, MICHAEL D THOMAS
PC C12N15/09, C12N1/21, C12N9/54//C07K14/325, (C12N1/21, C12R1:07),
CC C12N15/00
Process for producing polypeptide in bacillus cells FH Key
Location/Qualifiers
FT source 1. .185
/organism="Bacillus".
FEATURES
Source
1. .185
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 96.5%; Score 178.6; DB 6; Length 185;
Best Local Similarity 97.8%; Pred. No. 2.6e-33;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGT 60
Db 1 GGCCTTAAGGCGCTGCAATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGT 60
QY 61 CTGTCATCAGACAGGGTATTTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
Db 61 CTGTCATCAGACAGGGTATTTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGA 120
QY 121 ATAAAGGGGGTGCATTTATTTACTGATATGATATATAATTTGATATAGAAATGG 180
Db 121 ATAAAGGGGGTGTATTATTTACTGATATGATATATAATTTGATATAGAAATGG 180
QY 181 AGCTC 185
Db 181 AGCTC 185

RESULT 10
LOCUS 162 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 24 from patent US 5621089.
ACCESSION I40596
VERSION I40596.1 GI:2082888
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 162)
AUTHORS Sloma, A. P.; Outtrup, H.; Dammann, C. and Aaslyng, D. A.
TITLE Nucleic acid constructs for the production of a Bacillus alkaline protease
JOURNAL Patent: US 5621089-A 24 15-APR-1997;
FEATURES
Source
1. .162
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 84.1%; Score 155.6; DB 6; Length 162;
Best Local Similarity 97.5%; Pred. No. 1.1e-27;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 18 ATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGTGTATCAGACAGGT 77
Db 1 ATCGATTGTTTGAGAAAGAGAGACCAATAAAATACCTTGTGTATCAGACAGGT 60
QY 78 ATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGAATAGGGGGTTCACA 137
Db 61 ATTTTATGCTGCTCCAGACTGTCGGCTGTGTATATAAAGGAATAGGGGGTTCGTA 120
```

Qy	78	ATTTTATTGCTGTCACAGACTGCCGTGTGTAATAAAGAAGGAATAAAGGGGGTTGACA	137
Dd	61	ATTTTATTGCTGTCACAGACTGCCGTGTGTAATAAAGAAGGAATAAAGGGGGTTGTA	120
Qy	138	TATTTTACTGATATGATATATAATAATTTGTATAGAAAATG	179
Dd	121	TATTTTACTGATATGATATATAATAATTTGTATAGAAAATG	162
RESULT 13			
I56837			
LOCUS		162 bp	DNA
DEFINITION		Sequence 24 from patent US 5650326.	
ACCESSION		I56837	
VERSION		I56837.1	GI:2477250
KEYWORDS			
SOURCE		Unknown.	
ORGANISM		Unclassified.	
REFERENCE		1 (bases 1 to 162)	
AUTHORS		Sloma, A.P.; Outtrup, H.; Dambmann, C. and Aslyng, D. Anita.	
TITLE		Promoter element and signal peptide of a gene encoding a Bacillus alkaline protease and vectors comprising same	
JOURNAL		Patent: US 5650326-A 24 22-JUL-1997;	
FEATURES		Location/Qualifiers	
source		1..162	
ORIGIN		/organism="unknown"	
		/mol_type="unassigned DNA"	
Query Match		84.1%; Score 155.6; DB 6; Length 162;	
Best Local Similarity		97.5%; Pred. No. 1.1e-27;	
Matches 158; Conservative		0; Mismatches 4; Indels 0; Gaps 0;	
Qy	18	ATCCGATTGTTTGAGAAAGAGAGACCATAAAAAATCACCTGCTGTCATCAGACAGGCT	77
Dd	1	ATCCGATTGTTTGAGAAAGAGAGACCATAAAAAATCACCTGCTGTCATCAGACAGGCT	60
Qy	78	ATTTTATTGCTGTCACAGACTGCCGTGTGTAATAAAGAAGGAATAAAGGGGGTTGACA	137
Dd	61	ATTTTATTGCTGTCACAGACTGCCGTGTGTAATAAAGAAGGAATAAAGGGGGTTGTA	120
Qy	138	TATTTTACTGATATGATATATAATAATTTGTATAGAAAATG	179
Dd	121	TATTTTACTGATATGATATATAATAATTTGTATAGAAAATG	162
RESULT 14			
BD174625		249 bp	DNA
LOCUS			
DEFINITION		Modified promoter.	
ACCESSION		BD174625	
VERSION		BD174625.1	GI:29120315
KEYWORDS		JP 2002272466-A/14.	
SOURCE		Bacillus amyloliquefaciens	
ORGANISM		Bacillus amyloliquefaciens	
REFERENCE		Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.	
AUTHORS		Mizubuchi, H., Fushimi, N. and Miyoshi, S.	
TITLE		Modified promoter.	
JOURNAL		Patent: JP 2002272466-A 14 24-SEP-2002;	
COMMENT		OS SHOWA SANGYO CO LTD	
		SN Bacillus amyloliquefaciens	
		FN JP 2002272466-A/14	
		PD 24-SEP-2002	
		PF 15-MAR-2001	JP 2001074780
		PI HIROYUKI MIZUBUCHI, NAOKA FUSHIMI, SHINSUKE MIYOSHI PC	
		C12N15/09, C12N1/21, C12P21/02// (C12N1/21, C12R1:07), PC	
		C12R1:01,	
		PC (C12N1/21, C12R1:19), C12N15/00	
		CC Modified promoter	
FH Key		Location/Qualifiers	
FT source		1..249	

```

FT          /organism='Bacillus amyloliquefaciens'
FEATURES    Location/Qualifiers
   source    1..249
             /organism="Bacillus amyloliquefaciens"
             /mol_type="genomic DNA"
             /db_xref="taxon:1390"

ORIGIN
Query Match      84.1%; Score 155.6; DB 6; Length 249;
Best Local Similarity 97.5%; Pred. No. 9.7e-28;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 18 ATCGATTGTTGAGAAAGAGAGAGACCATATAAATACCTTCTGTCATCAGACAGGGT 77
    |||||
Db 73 ATCGATTGTTGAGAAAGAGAGAGACCATATAAATACCTTCTGTCATCAGACAGGGT 132
    |||||

Qy 78 ATTTTATGCTGCTCCAGACTGTCGCTGTGTAATAAAGGAATAAAGGGGGTTGACA 137
    |||||
Db 133 ATTTTATGCTGCTCCAGACTGTCGCTGTGTAATAAAGGAATAAAGGGGGTTGTTA 192
    |||||

Qy 138 TTATTTTACTGATGATATATAATTTGTATAGAAAATG 179
    |||||
Db 193 TTATTTTACTGATGATGTAATAATTTGTATAGAAAATG 234
    |||||

RESULT 15
BAMYL
LOCUS
DEFINITION Bacillus amyloliquefaciens gene fragment encoding alpha-amylase.
            (EC 3.2.1.1).
ACCESSION V00092
VERSION V00092.1 GI:39297
KEYWORDS amylase; amylase-alpha; signal peptide.
SOURCE Bacillus amyloliquefaciens
ORGANISM Bacillus amyloliquefaciens
REFERENCE 1 (bases 1 to 537)
AUTHORS Palva, I., Pettersson, R. F., Kalkkinen, N., Lehtovaara, P., Sarvas, M.,
Soderlund, H., Takkinen, K. and Kaariainen, L.
TITLE Nucleotide sequence of the promoter and NH2-terminal signal peptide
region of the alpha-amylase gene from Bacillus amyloliquefaciens
JOURNAL Gene 15 (1), 43-51 (1981)
MEDLINE 82051296
PUBMED 6170539
FEATURES
   source    Location/Qualifiers
   1..537
             /organism="Bacillus amyloliquefaciens"
             /mol_type="genomic DNA"
             /db_xref="taxon:1390"
             /note="unnamed protein product; amylase"
             /codon_start=1
             /transl_table=11
             /protein_id="CAA23430.1"
             /db_xref="GI:39298"
             /db_xref="GOA:P00692"
             /db_xref="SWISS-PROT:P00692"
             /translation="MIQKARTVSRFLVLMCTLLFVSLPTTKTSVAVNGILMQYFHWYT
PNDQHWKRLQNDAEHLSDIGITAVWIPAYKGLSOSDNGYGPYDLYLGEF"
             /sig_peptide 250..342
             /note="signal peptide"

ORIGIN
Query Match      84.1%; Score 155.6; DB 1; Length 537;
Best Local Similarity 97.5%; Pred. No. 8.3e-28;
Matches 158; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 18 ATCGATTGTTGAGAAAGAGAGAGACCATATAAATACCTTCTGTCATCAGACAGGGT 77
    |||||
Db 73 ATCGATTGTTGAGAAAGAGAGAGACCATATAAATACCTTCTGTCATCAGACAGGGT 132
    |||||

Qy 78 ATTTTATGCTGCTCCAGACTGTCGCTGTGTAATAAAGGAATAAAGGGGGTTGACA 137
    |||||
```

```

Db 133 ATTTTATGCTGCTCCAGACTGTCGCTGTGTAATAAAGGAATAAAGGGGGTTGTTA 192
Qy 138 TTATTTTACTGATGATATATAATTTGTATAGAAAATG 179
    |||||
Db 193 TTATTTTACTGATGATGTAATAATTTGTATAGAAAATG 234
    |||||
```

Search completed: March 19, 2004, 02:14:35
Job time : 1081.5 secs